

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 29, 2004, 12:58:43 ; Search time 85 Seconds
(without alignments)
1641.714 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389
Perfect score: 1938
Sequence: 1 MRSALLSAVALLRAQPPPP.....LRHNRRLPRFLFRNLSSL 389

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1938	100.0	566	3	AAY69183 Amino aci
2	1938	100.0	566	5	AAY78891 Mouse end
3	1477	76.2	560	2	AAR71294 Human gly
4	1477	76.2	560	3	AAY69184 Amino aci
5	1477	76.2	560	5	AAY78892 Human end
6	1477	76.2	560	7	ADP69095 Human MP5
7	607	31.3	581	5	ABBS3284 Human pol
8	607	31.3	581	6	ABUS2615 Human NOV
9	607	31.3	581	6	ABUS6706 Lung can
10	607	31.3	581	6	ABUS6544 Lung can
11	607	31.3	581	6	ABR43175 Human REM
12	607	31.3	581	7	ADCS3709 Human leu
13	607	31.3	581	7	ADCS3708 Rat leuci
14	607	31.3	581	7	ADCS3709 Novel pro
15	607	31.3	581	7	ADN39956 Cancer/an
16	607	31.3	790	7	ADN39092 Cancer/an
17	607	31.3	790	6	ADCS6793 Human GPC
18	583	30.1	578	6	ABP97224 Tumour-as
19	547	28.2	444	7	ADCS3708 Rat leuci
20	509.5	26.3	395	6	ADCS3709 Novel pro
21	503.5	26.0	570	8	ADCS36915 Human LRR
22	445	23.0	545	5	ABP64705 Human pro
23	445	23.0	545	6	ABJ26653 Human pro
24	445	23.0	545	6	ABUS4560 Human NOV
25	445	23.0	545	6	ABU12117 Human pro

26	442	22.8	545	6	ABUS4561 Human NOV
27	419.5	21.6	605	2	AAR85888 WD-40 dom
28	419.5	21.6	605	6	AAC23109 IGFALS 'h
29	419.5	21.6	605	7	ADD48887 Human Pro
30	419.5	21.6	605	7	ADD47835 Human Pro
31	419.5	21.6	605	7	ADDE1677 Human Pro
32	419.5	21.6	605	7	ADJ68826 Human hea
33	382.5	19.7	603	2	AAR85889 WD-40 dom
34	382.5	19.7	603	7	ADD48824 Rat Prote
35	382.5	19.7	603	7	ADD48885 Rat Prote
36	382.5	19.7	603	7	ADDE1675 Rat Prote
37	382.5	19.7	603	7	ADD47833 Rat Prote
38	379.5	19.6	928	8	AAO24336 Human sli
39	379.5	19.6	1122	8	AAO24335 Human sli
40	379.5	19.6	1461	8	AAO24327 Human sli
41	379.5	19.6	1508	2	AAV27141 Human sli
42	379.5	19.6	1508	2	AAW96706 Protein s
43	379.5	19.6	1508	2	AAV04138 Human sli
44	379.5	19.6	1531	7	ADG42633 Rat SLIT
45	379.5	19.6	1534	2	AAW46966 Amino aci

ALIGNMENTS

RESULT 1
AAY69183
ID AAY69183 standard; protein; 566 AA.
XX
AC AAY69183;
AC

DT 30-MAY-2000 (first entry)
XX
DE Amino acid sequence of murine glycoprotein V (GP V).
XX

KW Glycoprotein V; GP V; transgenic animal; platelet; thrombin substrate;
KW platelet function; haemostasis; platelet aggregation; anticoagulant;
KW thrombolytic; antiplatelet therapeutic drug.
XX
OS Mus sp.

XX
FH Key Location/Qualifiers
FT Misc-difference 566
FT /note= "encoded by TG"
XX
PN WO200008137-A2.
XX

PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017594.
XX
PR 04-AUG-1998; 98US-0109797P.
XX

PA (CORT-) COR THERAPEUTICS INC.
XX
PI Ramakrishnan V, Phillips DR;
XX
DR WPI; 2000-195570/17.
XX
N-PSDB; AAZ61216.
XX

PT Transgenic animals with a modified glycoprotein V gene used to identify
PT agents that modulate the effects of glycoprotein V.
XX
PS Example 1; Fig 2; 44pp; English.
XX

CC The present sequence represents a murine glycoprotein V (GP V). The
CC specification describes non-human transgenic animals which contain or
CC comprise a modified GP V gene. GP V is a platelet and endothelial cell
CC specific glycoprotein, and is a substrate for thrombin. Activation of
CC platelets by thrombin results in the loss of surface GP V. Platelets from
CC the transgenic animals can be used in a number of assays to identify
CC agents that modulate GP V function, or to assess the role of GP V in
CC platelet function. Such assays may help elucidate the extent to which GP

CC V is critical for normal haemostasis. The transgenic animals may also be
 CC used to identify agents that modulate platelet aggregation. Agents that
 CC can be tested include anticoagulants, thrombolytics, and antiplatelet
 CC therapeutic drugs
 XX
 SQ Sequence 566 AA;

Query Match 100.0%; Score 1938; DB 3; Length 566;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRSALLSAVALLRAQPPCPKTKCVVRDAAQCSCGGSVAHIAELGLPTNLTTHILLPRM 60
 DB 1 MRSALLSAVALLRAQPPCPKTKCVVRDAAQCSCGGSVAHIAELGLPTNLTTHILLPRM 60
 QY 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRPRALDK 120
 DB 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRPRALDK 120
 QY 121 MVLEQLFDHNALRDQNLFOQLRNQELGNQQLSFLPANLFSSRLRELKLDLSRN 180
 DB 121 MVLEQLFDHNALRDQNLFOQLRNQELGNQQLSFLPANLFSSRLRELKLDLSRN 180
 QY 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNQLTGALTTELRLERNHLSVAPGAFDR 240
 DB 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNQLTGALTTELRLERNHLSVAPGAFDR 240
 QY 241 LGNLSLTLTSGNLSLPPALFLHVSVSRILTFENPLEELPDVLFGEAGRELWNGT 300
 DB 241 LGNLSLTLTSGNLSLPPALFLHVSVSRILTFENPLEELPDVLFGEAGRELWNGT 300
 QY 301 HLSTLPAAPFNLSGLQTLGTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360
 DB 301 HLSTLPAAPFNLSGLQTLGTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360
 QY 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSSL 389
 DB 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSSL 389

RESULT 2

AAU7889;
 ID AAU78891 standard; protein; 566 AA.

AC AAU78891;
 XX
 XX
 DT 02-JUL-2002 (first entry)
 DE Mouse endothelial cell specific glycoprotein GP V.
 XX Endothelial cell specific glycoprotein; GP V; anticoagulant;
 XX platelet activation inhibitor; thrombin-induced activity;
 XX transgenic animal; anti-thrombotic agent; arterial thrombosis;
 XX atherosclerotic artery; acute myocardial infarction; stroke;
 XX ischaemic complication of cardiovascular disease;
 XX Bernard-Soulier syndrome; BSS; mouse.

XX Mus sp.

XX WO200217711-A2.

XX 07-MAR-2002.

XX 31-AUG-2001; 2001WO-US026936.

XX 31-AUG-2000; 2000US-0229047P.

XX 31-AUG-2000; 2000US-0230566P.

XX (CORT-) COR THERAPEUTICS INC.

XX Ramakrishnan V, Phillips D;

XX WPI; 2002-315494/35.

DR N-PSDB; ABK48274.

XX Identifying agent inhibiting thrombin-induced platelet activation, by
 PT administering test agent and proteolytically inactive thrombin to GP V
 PT null non-human transgenic animal and monitoring platelet aggregation.
 XX

PS Example 1; Fig 2; 64pp; English.

XX The invention describes a method of identifying an agent that inhibits
 CC thrombin-induced activity (e.g., platelet activation). Thrombin-induced
 CC activity can be modulated (inhibited) by endothelial cell specific
 CC glycoprotein GP V, by administering test agent and (proteolytically
 CC inactive) thrombin to a GP V null non-human transgenic animal (I), and
 CC monitoring aggregation of platelets in (I) to identify inhibition. The
 CC method is useful for identifying agents that inhibit thrombin-induced
 CC activity e.g. thrombin-induced platelet activation, and for screening
 CC anti-thrombotic agents. Inhibitors of thrombin-induced activity are
 CC useful for inhibiting arterial thrombosis in atherosclerotic arteries
 CC e.g. acute myocardial infarction and stroke, inducing ischaemic
 CC complications of cardiovascular disease and platelet activation in a
 CC subject and may be useful for treatment of Bernard-Soulier syndrome
 CC (BSS). This is the amino acid sequence of the murine endothelial cell
 CC specific glycoprotein GP V isolated in the invention

XX Sequence 566 AA;

Query Match 100.0%; Score 1938; DB 5; Length 566;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSALLSAVALLRAQPPCPKTKCVVRDAAQCSCGGSVAHIAELGLPTNLTTHILLPRM 60
 DB 1 MRSALLSAVALLRAQPPCPKTKCVVRDAAQCSCGGSVAHIAELGLPTNLTTHILLPRM 60
 QY 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRPRALDK 120
 DB 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRPRALDK 120
 QY 121 MVLEQLFDHNALRDQNLFOQLRNQELGNQQLSFLPANLFSSRLRELKLDLSRN 180
 DB 121 MVLEQLFDHNALRDQNLFOQLRNQELGNQQLSFLPANLFSSRLRELKLDLSRN 180
 QY 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNQLTGALTTELRLERNHLSVAPGAFDR 240
 DB 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNQLTGALTTELRLERNHLSVAPGAFDR 240
 QY 241 LGNLSLTLTSGNLSLPPALFLHVSVSRILTFENPLEELPDVLFGEAGRELWNGT 300
 DB 241 LGNLSLTLTSGNLSLPPALFLHVSVSRILTFENPLEELPDVLFGEAGRELWNGT 300
 QY 301 HLSTLPAAPFNLSGLQTLGTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360
 DB 301 HLSTLPAAPFNLSGLQTLGTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360
 QY 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSSL 389
 DB 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSSL 389

RESULT 3

AAR71294

ID AAR71294 standard; protein; 560 AA.

XX AAR71294;

XX 25-MAR-2003 (revised)

DT 18-AUG-1995 (first entry)

XX Human glycoprotein V.

DE Glycoprotein V; GPV; platelet.

XX Homo sapiens.

OS

```

XX PH Key Location/Qualifiers
FT Peptide 1..16
FT Modified-site /label= Sig_peptide
FT Modified-site 51
FT Modified-site /label= N-glycosylation_site
FT Modified-site 181
FT Modified-site /label= N-glycosylation_site
FT Modified-site 244
FT Modified-site /label= N-glycosylation_site
FT Modified-site 267
FT Modified-site /note= "N-glycosylation_site"
FT Modified-site 298
FT Modified-site /label= N-glycosylation_site
FT Modified-site 312
FT Modified-site /label= N-glycosylation_site
FT Modified-site 385
FT Modified-site /label= N-glycosylation_site
FT Cleavage-site 476..477
FT Modified-site /note= "putative thrombin cleavage site"
FT Modified-site 499
FT Domain /label= N-glycosylation_site
FT Domain 520..544
FT Domain /note= "putative transmembrane domain"
XX W09502054-A2.
XX XX
XX 19-JAN-1995.
XX 07-JUL-1994; 94WO-US007644.
XX 09-JUL-1993; 93US-00089455.
XX 03-DEC-1993; 93US-00162599.
XX 10-FEB-1994; 94US-00195006.
XX (CORT-) COR THERAPEUTICS INC.
XX Lanza F, Phillips DR, Cazenave J;
XX WPI; 1995-066899/09.
XX N-PSDB; AAQ85594.
XX Platelet glycoprotein V gene - useful for producing glycoprotein V (GPV)
XX and variants and generating antibodies to GPV.
XX Disclosure; Page 45-50; 82pp; English.
XX Genomic clones were isolated from a human fibroblast library in lambda
XX Fix using a 748 bp 32P-labeled glycoprotein V (GPV) cDNA probe. Exon-
XX containing fragments from positive clones were subcloned and sequenced.
XX The full sequence of the human GPV gene is given in AAQ85594. (Updated on
XX 25-MAR-2003 to correct FN field.)
XX SQ Sequence 560 AA;

Query Match 76.2%; Score 1477; DB 2; Length 560;
Best Local Similarity 76.1%; Pred. No. 1e-127;
Matches 295; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVALLRAQPPCPKTKCVVDRDAQCSGGSVVAHIAEGLPTNTHILLFRM 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 MLRGTLCAVLGLRAQPPCPACKCFRDAQCSGGDVARIASALGLPTNTHILLFCM 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 DQGLIRHNSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLKTLRTRNKISBLPRAIDK 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GRGVLSQSFSGMTVLQRLMISHSISAVAPGFSDLIKTKLRLSRNKITHLPALLDX 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 MVLLQFLFDHNAIRLDQNLFOQLRNQLBELGNQQLSFLPANFSSRLKLDLSRN 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 MVLLQFLFDHNAIRLDQNMFKVLNQLBELANQQLDFLPASUFTNLKLLDLSGN 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 NLTHLPKGLGKGLGKLLYSNQLTSVDGSLLSNLGALTELRERNHRSVAPGADR 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 181 NLTHLPKGLGKGLGKLLYSNQLTSVDGSLLSNLGALTELRERNHRSVAPGADR 240
QY 241 LCNLSLTLSGNLLESLPPALFLHVSVSRLTLFENPLEELPDVLFGSMAGRLWLNGT 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 LFNLSLTLSRNHLAFLPSALFLSHNLTLLTLFENPLAELPGVLFGSMGGIQLWLNR 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFGQLRELRVLALHTNALAELEDDALR 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 QLRTLPAAAFRNLSRLRYLVGVTLSPLRSALPQGAFOGLGELQVLAHNSGTLALPDGLLR 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GLGHLRQVSLRHRNRLRALPRTLFRNLSSL 389
Db 361 GLGKLRQVSLRHRNRLRALPRTLFRNLSSL 389

RESULT 4
AAY69184
ID AAY69184 standard; protein; 560 AA.
XX
AC AAY69184;
XX
DT 30-MAY-2000 (first entry)
XX
DE Amino acid sequence of a human glycoprotein V (GP V).
XX
KW Glycoprotein V; GP V; transgenic animal; platelet; thrombin substrate;
KW platelet function; haemostasis; platelet aggregation; anticoagulant;
KW thrombolytic; antiplatelet therapeutic drug.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 163 /note= "encoded by CCC"
FT Misc-difference 485 /note= "encoded by CAC"
FT Misc-difference 490 /note= "encoded by GDC"
FT Misc-difference 547 /note= "encoded by CGC"
FT FT FT FT
XX WC200008137-A2.
XX
XX 17-FEB-2000.
XX
XX 04-AUG-1999; 99WO-US017594.
XX
XX 04-AUG-1998; 98US-0109797P.
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Ramakrishnan V, Phillips DR;
XX
XX WPI; 2000-195570/17.
XX N-PSDB; AAZ61217.
XX Transgenic animals with a modified glycoprotein V gene used to identify
XX agents that modulate the effects of glycoprotein V.
XX Example 1; Fig 4; 44pp; English.
XX The present sequence represents a human glycoprotein V (GP V). The
XX specification describes non-human transgenic animals which contain or
XX comprise a modified GP V gene. GP V is a platelet and endothelial cell
XX specific glycoprotein, and is a substrate for thrombin. Activation of
XX platelets by thrombin results in the loss of surface GP V. Platelets from
XX the transgenic animals can be used in a number of assays to identify
XX agents that modulate GP V function, or to assess the role of GP V in
XX platelet function. Such assays may help elucidate the extent to which GP
XX V is critical for normal haemostasis. The transgenic animals may also be
XX used to identify agents that modulate platelet aggregation. Agents that
XX can be tested include anticoagulants, thrombolytics, and antiplatelet
XX therapeutic drugs

```

```

XX SQ Sequence 560 AA;
Query Match 76.2%; Score 1477; DB 3; Length 560;
Best Local Similarity 76.1%; Pred. No. 1e-127;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPKTCVVDAAQCGSGSVAHIAELGLPTNLTHILLFRM 60
DB 1 MLRGTLCAVLGLLRAQPPCPACKCVFRDAAQCGGVDARISALGLPTNLTHILLFGM 60
QY 61 DOGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRLPAILDK 120
DB 61 GRGVLSQSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRLPAILDK 120
QY 121 MVLLQQLFDHNAERLDQNLFOQLRNQELGNQNLSPANLFSRLKLLDLSRN 180
DB 121 MVLLQQLFDHNAERLDQNLFOQLRNQELGNQNLSPANLFSRLKLLDLSRN 180
QY 181 NLTHLPKGLGAGQVKLEKLLYSNQLTSVDSGLSNLQALTELRLERNHLSVAPGAFDR 240
DB 181 NLTHLPKGLGAGQVKLEKLLYSNQLTSVDSGLSNLQALTELRLERNHLSVAPGAFDR 240
QY 241 LGLHRLQVSLRHNRLRALPRTLFRNLSSL 389
DB 361 GLGKLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 5
AAU78892
ID AAU78892 standard; protein; 560 AA.
AC AAU78892;
XX
XX DT 02-JUL-2002 (first entry)
XX
DE Human endothelial cell specific glycoprotein GP V.
KW Endothelial cell specific glycoprotein; GP V; anticoagulant;
KW platelet activation inhibitor; thrombin-induced activity;
KW transgenic animal; anti-thrombotic agent; arterial thrombosis;
KW atherosclerotic artery; acute myocardial infarction; stroke;
KW ischaemic complication of cardiovascular disease;
KW Bernard-Soulier syndrome; BSS; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 162 /note= "Encoded by CCC"
FT Misc-difference 419 /note= "Encoded by GCG"
FT Misc-difference 485 /note= "Encoded by CAC"
FT Misc-difference 490 /note= "Encoded by GDC"
XX
XX WO200217711-A2.
XX
XX PD 07-MAR-2002.
XX
XX PF 31-AUG-2001; 2001WO-US026936.
XX
XX PR 31-AUG-2000; 2000US-0229047P.
XX
XX PR 31-AUG-2000; 2000US-0230566P.
XX

```

(CORT-) COR THERAPEUTICS INC.
Ramakrishnan V, Phillips D;
WPI; 2002-315494/35.
DR N-PSDB; ABX48275.

Identifying agent inhibiting thrombin-induced platelet activation, by administering test agent and proteolytically inactive thrombin to GP V null non-human transgenic animal and monitoring platelet aggregation.

Disclosure; Fig 4; 64pp; English.

The invention describes a method of identifying an agent that inhibits thrombin-induced activity (e.g., platelet activation). Thrombin-induced activity can be modulated (inhibited) by endothelial cell specific glycoprotein GP V, by administering test agent and (proteolytically inactive) thrombin to a GP V null non-human transgenic animal (I), and monitoring aggregation of platelets in (I) to identify inhibition. The method is useful for identifying agents that inhibit thrombin-induced activity e.g. thrombin-induced platelet activation, and for screening anti-thrombotic agents. Inhibitors of thrombin-induced activity are useful for inhibiting arterial thrombosis in atherosclerotic arteries e.g. acute myocardial infarction and stroke, inducing ischaemic complications of cardiovascular disease and platelet activation in a subject and may be useful for treatment of Bernard-Soulier syndrome (BSS). This is the amino acid sequence of the human endothelial cell specific glycoprotein GP V isolated in the invention

XX SQ Sequence 560 AA;

Query Match 76.2%; Score 1477; DB 5; Length 560;
Best Local Similarity 76.1%; Pred. No. 1e-127;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPKTCVVDAAQCGSGSVAHIAELGLPTNLTHILLFRM 60
DB 1 MLRGTLCAVLGLLRAQPPCPACKCVFRDAAQCGGVDARISALGLPTNLTHILLFGM 60
QY 61 DOGILRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRLPAILDK 120
DB 61 GRGVLSQSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTIRLNRKISRLPAILDK 120
QY 121 MVLLQQLFDHNAERLDQNLFOQLRNQELGNQNLSPANLFSRLKLLDLSRN 180
DB 121 MVLLQQLFDHNAERLDQNLFOQLRNQELGNQNLSPANLFSRLKLLDLSRN 180
QY 181 NLTHLPKGLGAGQVKLEKLLYSNQLTSVDSGLSNLQALTELRLERNHLSVAPGAFDR 240
DB 181 NLTHLPKGLGAGQVKLEKLLYSNQLTSVDSGLSNLQALTELRLERNHLSVAPGAFDR 240
QY 241 LGLHRLQVSLRHNRLRALPRTLFRNLSSL 389
DB 361 GLGKLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 6
ADF69095
ID ADF69095 standard; protein; 560 AA.
XX
XX AC ADF69095;
XX
XX DT 12-FEB-2004 (first entry)
XX

DE Human MP53 protein sequence SEQ ID NO:65.
XX
KW p53 pathway modulating agent; MP53; p53 modulator; cytostatic;
KW gene therapy; cancer; human.
XX
OS Homo sapiens.
XX
FN WO2003083047-A2.
XX
PD 09-OCT-2003.
XX
PF 28-FEB-2003; 2003WO-US06025.
XX
PR 01-MAR-2002; 2002US-0361196P.
XX
FA (EXEL-) EXELIXIS INC.
XX
PI Belvin M, Francis-Lang H, Friedman L, Plowman GD, Heuer TS, Li D;
PI Funke RP;
XX
XX WPI: 2003-812540/76.
DR N-PSDB; ADF69151.
XX
XX Identifying a candidate p53 pathway modulating agent for treating e.g.,
PT cancer by contacting an assay system comprising a MP53 polypeptide or
PT nucleic acid with a test agent and detecting a test agent-biased
PT activity.
XX
XX Example; SEQ ID NO 65; 406pp; English.
PS
XX
CC The present invention describes a method for identifying a candidate p53
CC pathway modulating agent, which comprises: (a) providing an assay system
CC comprising a MP53 (modulator of p53) polypeptide or nucleic acid or its
CC fragment or derivative; (b) contacting the assay system with a test agent
CC under conditions where the system provides a reference activity except in
CC the presence of the test agent; and (c) detecting a test agent-biased
CC activity, where a difference between the test agent-biased activity and
CC the reference activity identifies the test agent as a candidate p53
CC pathway modulating agent. Also described: (1) modulating the p53 pathway
CC of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3)
CC diagnosing a disease in a patient. MP53 has cytostatic activity, and can
CC be used in gene therapy. The method is useful for identifying a candidate
CC p53 pathway modulating agent for preparing a composition for diagnosing
CC or treating e.g., cancer. The present sequence represents a human MP53
CC protein, which is used in the exemplification of the present invention.
XX
SQ Sequence 560 AA;

Query Match 76.2%; Score 1477; DB 7; Length 560;
Best Local Similarity 76.1%; Pred. No. 1e-127;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALRAQPPFPCKTKCVRDAAQCSGGSVAHTAELGLPTNLTHILLRPM 60
DB 1 MLRGTLCAVLGLLRAQPPFPCKCVFDDAAQCSGSDVARLSALGLPTNLTHILLFGM 60
QY 61 DQILRNHSGMTVLORLMLSDSHISAIDPGTFNDLVKLTURLTRNKISRLPAILDK 120
DB 61 GRGVLOSQSGMTVLQRLMISDSHISAVAPGTFSDILKLTURLSRNKITHLPGLLDK 120
QY 121 MVLEQLFDHNLARDLQNLQRLNQLGHNQCLSPANLPSLSRELKLLDLSN 180
DB 121 MVLEQLFDHNLARGDQNNFQKLVNLQELANQQLDPLASLFTNLENKLLDLSN 180
QY 181 NLTHLPKGLGAQVKLEKLLYLNQLTSDVSLGSLNGLALTEURLERNLRSVAPGAFDR 240
DB 181 NLTHLPKGLGAQAKLERLLHNSRLVSLDSGLNSLGLALTELPQPHNRHRSIAPGAFDR 240
QY 241 LGNLSLTLGSLNLESLPPALFHVSVSELTITFENPLELPVLGEGVAGRELWNGT 300
DB 241 LFNLSLTLGRNHLAFPLSALFHSNMLTLTLFENPLAELPGLVFGEGGQELWNGT 300
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRVIALHTNALAELRDDALR 360

Db 301 QLRTLPAFAAFRNLSRLRYLGVLTSPRLSALPQAGLQGLQVLAHNSGLTALPDGLLR 360
QY 361 GLGHLRQVSLRNLRALPRTLFRNLSSL 389
Db 361 GLGKLRQVSLRRNRALPFRNLSSL 389

RESULT 7
ABBS3264
ID ABBS3264 standard; protein; 581 AA.
XX
AC ABBS3264;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human polypeptide #4.
XX
KW Human; nootropic; neuroprotective; anticonvulsant; antidepressant;
KW neuroleptic; tranquiliser; antiarrhythmic; cardiant; antiasthmatic;
KW antiinflammatory; antilipemic; hepatotropic; virucide; antidiabetic;
KW nephrotropic; anorectic; cytostatic; vaccine; neurological disease;
KW cardiovascular disease; respiratory disease; liver disease;
KW renal disease; skeletal muscle disease; gastrointestinal disease;
KW placental disease; testicular cancer; male fertility; pancreatic disease.
XX
OS Homo sapiens.
XX
XX WO200181363-A1.
XX
PD 01-NOV-2001.
XX
XX 26-APR-2001; 2001WO-US013360.
XX
PF 27-APR-2000; 2000US-0199963P.
PR 11-MAY-2000; 2000US-0203336P.
PR 25-MAY-2000; 2000US-0207087P.
PR 26-MAY-2000; 2000US-0207546P.
XX
XX (SMIX) SMITHKLINE BEECHAM CORP.
XX (SMIX) SMITHKLINE BEECHAM PLC.
XX
PI Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;
PI Lai Y, Xie Q;
XX
XX WPI: 2002-041392/05.
DR N-PSDB; ABA90329.
XX
XX Novel polypeptides and polynucleotides useful as a vaccine for preventing
PT and treating diseases associated the polypeptide, e.g. Alzheimer's
PT disease, dyslipidemia, obesity, diabetes, infertility, asthma, amnesias.
XX
XX Claim 1; Page 70-71; 116pp; English.
PS
XX
CC The invention relates to an isolated polypeptide comprising a 277, 480,
CC 583, 581, 628, 428, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,
CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,
CC 784, 252, 593, 472, 607, 781, 640, 686 or 154 amino acid sequence as
CC given in the specification. The polypeptides, modulators of the
CC polypeptides and antibodies against the polypeptides are useful for
CC treating diseases such as neurological and psychiatric diseases including
CC Alzheimer's, parapsychic nuclear palsy, Huntington's disease, myotonic
CC dystrophy, anorexia and depression; cardiovascular diseases including
CC congestive heart failure, Hodgkin's disease and myocardial infarction;
CC respiratory diseases including asthma, chronic obstructive pulmonary
CC disease, cystic fibrosis and adult respiratory distress syndrome; liver
CC disease including hypercholesterolemia, cirrhosis, viral and nonviral
CC hepatitis, Type II diabetes mellitus, and impaired glucose tolerance;
CC renal disease including renal failure, acute tubular necrosis and
CC glomerulonephritis; skeletal muscle diseases including Eulenburg's
CC disease, hypoglycaemia and obesity; gastrointestinal diseases including
CC myotonia congenita and intestinal obstruction; lymph diseases including
CC lymphagiectasia; diseases of placenta including choriocarcinoma; diseases

CC proteins given in ABR43169 to ABR43204 and designated REMAP-1 to REMAP-36
CC (1). (1) have cytostatic, antiarteriosclerotic, anticonvulsant,
CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
CC antiinflammatory and thyromimetic activities, and can be used in gene
CC therapy. The REMAP polypeptides and polynucleotides are useful in
CC diagnosing, treating and preventing diseases or conditions associated
CC with the decreased expression or overexpression of REMAP, such as cell
CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome).
CC disorders, or infections. They are also useful in assessing the effects
CC of exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of REMAP. The REMAPs or their fragments are useful in screening
CC compounds for effectiveness as agonist or antagonist of the polypeptide,
CC or in altering the expression of the target polynucleotide and compounds
CC that specifically bind to or modulate the activity of the polypeptide
XX

SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;
Best Local Similarity 37.7%; Pred. No. 4.4e-47;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVDRDAQCSCGSAHIAELGLPTNLTHILLFRMDQGLRNHSPFGMTVLQRLM 80
DB 25 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQTLNTHITELNESPFNLISALIALR 83
QY 81 LSDSHISAIDPGTFNDLVKLTIRLTKNKSRLPRAILDQVLLLEQLFLDHNALRDLDQN 140
DB 84 IEKNELSRITPGAFNGLSGRYLSLANNKLVLPGLFQGLDLSLESLSSNQLLQIPA 143
QY 141 LFQQLRNQLQELGNQNLSPANLFSLSRELKLDLSRNMLTHLPKGLGAQVKLEKL 200
DB 144 HFSQCSNLKELQHLGNHLEYIPDGAFDHLVGLTKNLGNKSLTHISPRVFOHGLNQLVLR 203
QY 201 LYSNQLTSVDSGLLSNLCALTELRNHLRSVAPGAFDLGNLSLITLSCNLESPPA 260
DB 204 LYENRLTDIPMGTFDGLVNLQELALQOQIGLSFGLFHNHNLQRLYLSNNHISQLPPS 263
QY 261 LFLHVSSVRLTLFENPLELPDVLFGEMAGRELWNGTHLSTLPAAAFNLSGLQTLG 320
DB 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMNRLWYDNHISLSDPNVFSNLRQLQVLI 323
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALRGHLRQVSLRHNRRLALPR 380
DB 324 LSRN-QISFTSPGAFNGLTELRSLHTNALQDLGDNVFRMLANQLNISLQNNRLQLPG 382
QY 381 TLFNRLSSL 389
DB 383 NIFANVNGL 391

RESULT 12

ADC83709
ID ADC83709 standard; protein; 581 AA.

XX AC ADC83709;

XX 01-JAN-2004 (first entry)

XX DE Human leucine rich repeat (LRR) Lib protein.

XX KW Human; Lib; chromosome 3q29; leucine-rich repeat; LRR;
KW transmembrane domain; adhesion; cell-cell; cell-extracellular matrix;
KW astroglial cell; beta-amyloid processing; Alzheimer's disease;
KW neuron death; drug screening; cell adhesion-related disease; neurotropic;
KW neuroprotective; gene therapy.

OS Homo sapiens.

XX JP2003164290-A.

XX 10-JUN-2003.

RESULT 13

ADE08191

ID ADE08191 standard; protein; 581 AA.

XX 30-NOV-2001; 2001JP-00367093.
XX 30-NOV-2001; 2001JP-00367093.
XX (DAUC) DAIICHI PHARM CO LTD.
XX WPI; 2003-630582/60.
XX N-PSDB; ADC83707.
XX Novel gene encoding a protein having cell-cell adhesion or cell-
XX extracellular matrix adhesion regulatory property useful for diagnosing
XX or treating neuron death, and Alzheimer's disease.
XX Claim 3; Page 14-16; 20pp; Japanese.

XX The invention relates to rat and human proteins designated Lib (ADC83708-
XX ADC83709), and cDNAs encoding them (ADC83706-ADC83707). The Lib proteins
XX contain a leucine-rich repeat (LRR) and a transmembrane domain, and are
XX involved in regulation of cell-cell adhesion or cell-extracellular matrix
XX adhesion. They are expressed in astroglial cells and their expression is
XX induced by beta-amyloid processing, indicating that they may be useful in
XX the diagnosis, treatment or prevention of Alzheimer's disease. The
XX invention also encompasses polynucleotides with at least 70% homology to
XX the Lib polynucleotide, recombinant vectors and host cells comprising a
XX Lib polynucleotide, the recombinant expression of Lib proteins, an
XX antibody against a Lib protein, and methods of screening for modulators
XX of cell adhesion. Lib proteins and their encoding nucleic acids may be
XX used in the diagnosis of Alzheimer's disease or neuron death, and also to
XX screen for modulators of its activity. Such compounds are useful for
XX treating or preventing Alzheimer's disease-related neuron death or a
XX disease involving cell-cell adhesion or cell-extracellular matrix
XX adhesion. The present sequence represents human Lib protein. The human
XX Lib gene encoding it is located on chromosome 3q29.

XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;
Best Local Similarity 37.7%; Pred. No. 4.4e-47;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVDRDAQCSCGSAHIAELGLPTNLTHILLFRMDQGLRNHSPFGMTVLQRLM 80
DB 25 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQTLNTHITELNESPFNLISALIALR 83
QY 81 LSDSHISAIDPGTFNDLVKLTIRLTKNKSRLPRAILDQVLLLEQLFLDHNALRDLDQN 140
DB 84 IEKNELSRITPGAFNGLSGRYLSLANNKLVLPGLFQGLDLSLESLSSNQLLQIPA 143
QY 141 LFQQLRNQLQELGNQNLSPANLFSLSRELKLDLSRNMLTHLPKGLGAQVKLEKL 200
DB 144 HFSQCSNLKELQHLGNHLEYIPDGAFDHLVGLTKNLGNKSLTHISPRVFOHGLNQLVLR 203
QY 201 LYSNQLTSVDSGLLSNLCALTELRNHLRSVAPGAFDLGNLSLITLSCNLESPPA 260
DB 204 LYENRLTDIPMGTFDGLVNLQELALQOQIGLSFGLFHNHNLQRLYLSNNHISQLPPS 263
QY 261 LFLHVSSVRLTLFENPLELPDVLFGEMAGRELWNGTHLSTLPAAAFNLSGLQTLG 320
DB 264 IFMQLPQLNRLTLFGNSLKELSLGIFGPMNRLWYDNHISLSDPNVFSNLRQLQVLI 323
QY 321 LTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALRGHLRQVSLRHNRRLALPR 380
DB 324 LSRN-QISFTSPGAFNGLTELRSLHTNALQDLGDNVFRMLANQLNISLQNNRLQLPG 382
QY 381 TLFNRLSSL 389
DB 383 NIFANVNGL 391

RESULT 13

ADE08191

ID ADE08191 standard; protein; 581 AA.

CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;
 Best Local Similarity 37.7%; Pred. No. 4.4e-47;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVVRDAACSGGSAHIAELGPTNLTHTLLFRMDQGLRNHSPSGMTVLQRLM 80
 Db 25 CPSECTCSRAQSVCTGARIVAV-PTPLPWAMSLQINHTITELNESPFLNISALIALR 83
 QY 81 LSDSHISAIDPGTNDLVKLTLETRNKISRLPRAILDKNVLEQLFDHNAIRDQDN 140
 Db 84 IENKELSRITPGAFNLGSLRYLSLANNKLQVLPIGLFQGLDLSLESLLSSNQLQIQPA 143
 QY 141 LFQQLRNLOELGLNQNLSPFLPANSLSRELKLLDLSRNNLTPLPKGLIGAQVKLEKLL 200
 Db 144 HFSQCSNLKELQHLGNHLEYIPDGAFDHLVGLTKNLGNKSLTHISPRVFOHNLGNLQVLR 203
 QY 201 LYSNQLTSVPSGLLSNLTGALTELRNHLRSVAPGAFDRGLNLSLTLSCNLLSLPPA 260
 Db 204 LYENRLTDIPMGTDGLVNLQELAQOQIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263
 QY 261 LELHVSSVRLTLPENLEELPDVLFCEMAGLRELWNGHTLSTPLPAAFRNLGSGQLTG 320
 Db 264 IFMQLPOLNRLTFGNSKLSLGIQFGPMNRELWLDYDNHISLSDPNVFSNRLQVLI 323
 QY 321 LTRNPRLSALPRGVQFGLRELRLVLALHTNALAEIRDALRGHLRQVSRNHLRALPR 380
 Db 324 LSRN-QTSFISPGAFNGLTRELRLSLHTNALQDLQDGNVFRMLANQLNISLQNNRLQLPG 382
 QY 381 TLFNLSLL 389
 Db 383 NIFANVGL 391

RESULT 15
 ADN39092

ID ADN39092 standard; protein; 581 AA.

AC ADN39092;

DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:410.

XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; Gene therapy; vaccine.

XX Homo sapiens.

OS W02003042661-A2.

PN

XX

22-MAY-2003.

13-NOV-2002; 2002WO-US036810.

13-NOV-2001; 2001US-0350666P.

21-NOV-2001; 2001US-0332464P.

29-NOV-2001; 2001US-0334393P.

03-DEC-2001; 2001US-0335394P.

14-DEC-2001; 2001US-0340376P.

08-JAN-2002; 2002US-0342111P.

10-JAN-2002; 2002US-0347349P.

08-FEB-2002; 2002US-0355250P.

13-FEB-2002; 2002US-0356714P.

20-FEB-2002; 2002US-0359077P.

29-MAR-2002; 2002US-0368809P.

04-APR-2002; 2002US-0370110P.

12-APR-2002; 2002US-037246P.

05-JUN-2002; 2002US-0386614P.

16-JUL-2002; 2002US-0396839P.

22-JUL-2002; 2002US-039775P.

22-JUL-2002; 2002US-0397845P.

09-SEP-2002; 2002US-0409450P.

(EOSB-) EOS BIOTECHNOLOGY INC.

XX

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;

XX WPI; 2003-468649/44.

DR N-PSDB; ADN39091.

XX

PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.

XX

PS Claim 12; SEQ ID NO 410; 1385pp; English.

XX

CC The invention relates to nucleic acids and proteins (ADN39693-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;

Best Local Similarity 37.7%; Pred. No. 4.4e-47;

Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVVRDAACSGGSAHIAELGPTNLTHTLLFRMDQGLRNHSPSGMTVLQRLM 80

Db 25 CPSECTCSRAQSVCTGARIVAV-PTPLPWAMSLQINHTITELNESPFLNISALIALR 83

QY 81 LSDSHISAIDPGTNDLVKLTLETRNKISRLPRAILDKNVLEQLFDHNAIRDQDN 140

Db 84 IENKELSRITPGAFNLGSLRYLSLANNKLQVLPIGLFQGLDLSLESLLSSNQLQIQPA 143

QY 141 LFQQLRNLOELGLNQNLSPFLPANSLSRELKLLDLSRNNLTPLPKGLIGAQVKLEKLL 200

Db 144 HFSQCSNLKELQHLGNHLEYIPDGAFDHLVGLTKNLGNKSLTHISPRVFOHNLGNLQVLR 203

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:11:04 ; Search time 93 Seconds

(without alignments)
1483.616 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389

Perfect score: 1938

Sequence: 1 MRSALLSAVALLRAQPPF.....LHNRRLALPRLFRNLSSL 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1938	100.0	566	10	US-09-775-803-12
2	1938	100.0	567	15	US-10-037-417-105
3	1923	99.2	567	15	US-10-037-417-106
4	1749	90.2	567	15	US-10-080-334-228
5	1477	76.2	560	9	US-09-782-980-61
6	1477	76.2	560	10	US-09-775-803-14
7	1477	76.2	560	14	US-10-213-499-2
8	1477	76.2	560	16	US-10-806-018-61
9	1201	62.0	368	15	US-10-312-311-21
10	607	31.3	581	14	US-10-298-027-410
11	607	31.3	581	14	US-10-298-027-410
12	607	31.3	581	15	US-10-298-027-410
13	607	31.3	581	15	US-10-080-334-76
					Sequence 225, App

14	607	31.3	581	15	US-10-358-951-43	Sequence 43, Appl
15	607	31.3	787	15	US-10-080-334-224	Sequence 224, App
16	607	31.3	790	14	US-10-017-161-1558	Sequence 1558, App
17	607	31.3	790	14	US-10-292-798-1246	Sequence 1246, App
18	607	31.3	819	14	US-10-241-220-106	Sequence 106, App
19	607	31.3	819	17	US-10-872-972-106	Sequence 106, App
20	583	20.1	578	15	US-10-080-334-226	Sequence 226, App
21	541.5	27.9	391	15	US-10-080-334-227	Sequence 227, App
22	503.5	26.0	570	14	US-10-424-233-14	Sequence 14, Appl
23	445	23.0	545	15	US-10-114-270-38	Sequence 38, Appl
24	442	22.8	545	15	US-10-114-270-40	Sequence 40, Appl
25	419.5	21.6	457	15	US-10-312-311-22	Sequence 22, Appl
26	419.5	21.6	605	16	US-10-408-765A-632	Sequence 632, App
27	416.5	21.5	605	16	US-10-706-791-21	Sequence 21, Appl
28	413.5	21.3	510	15	US-10-312-311-20	Sequence 20, Appl
29	413.5	21.3	605	9	US-09-782-980-62	Sequence 62, Appl
30	413.5	21.3	605	16	US-10-806-018-62	Sequence 62, Appl
31	379.5	19.6	928	16	US-10-442-658-28	Sequence 28, Appl
32	379.5	19.6	1122	16	US-10-442-658-27	Sequence 27, Appl
33	379.5	19.6	1461	16	US-10-442-658-7	Sequence 7, Appl
34	373.5	19.6	1531	10	US-09-970-944-31	Sequence 31, Appl
35	373.5	19.6	1534	10	US-09-970-944-30	Sequence 30, Appl
36	373.5	19.6	1534	16	US-10-442-658-6	Sequence 6, Appl
37	373.5	19.4	1531	10	US-09-970-944-28	Sequence 28, Appl
38	369.5	19.1	796	14	US-10-028-392-5	Sequence 5, Appl
39	369.5	19.1	1529	10	US-09-866-050A-396	Sequence 396, App
40	367.5	19.0	907	14	US-10-271-078-10	Sequence 10, Appl
41	366.5	18.9	1348	15	US-10-080-334-234	Sequence 234, App
42	366	18.9	467	15	US-10-343-348-36	Sequence 36, Appl
43	362	18.7	421	15	US-10-072-012-294	Sequence 294, App
44	362	18.7	1523	10	US-09-954-342-45	Sequence 45, Appl
45	361.5	18.7	1529	10	US-09-766-511B-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-775-803-12
; Sequence 12, Application US/09775803
; Publication No. US20030167487A1
; GENERAL INFORMATION:
; APPLICANT: COR Therapeutics, Inc.
; APPLICANT: Ramakrishnan, Vanitha
; APPLICANT: Phillips, David
; TITLE OF INVENTION: Transgenic Animals Having a Modified Glycoprotein V
; FILE OF INVENTION: Trans and Methods for Their Use
; FILE REFERENCE: 44481-5044-US
; CURRENT APPLICATION NUMBER: US/09/775,803
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: PCT/US99/17594
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 60/109,797
; PRIOR FILING DATE: 1998-08-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-775-803-12

Query Match 100.0%; Score 1938; DB 10; Length 566;
Best Local Similarity 100.0%; Pred. No. 4.6e-156;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSALLSAVALLRAQPPFPCPKCKCVRDAAQCGSVAHIAELGLPTNLTHILFRM 60
Db 1 MRSALLSAVALLRAQPPFPCPKCKCVRDAAQCGSVAHIAELGLPTNLTHILFRM 60
QY 61 DQGIIRNHSFGMTVQLRLMLSDSHISAIDPGTNDLVKLTURLTRNKISRLPRAILDK 120
Db 61 DQGIIRNHSFGMTVQLRLMLSDSHISAIDPGTNDLVKLTURLTRNKISRLPRAILDK 120

QY 121 MVLLEQLFDHNAFLDQNLFOQLRNQLGELGNQNLQSLFPLANLFSLSRELKLLDLSRN 180
DB 121 MVLLEQLFDHNAFLDQNLFOQLRNQLGELGNQNLQSLFPLANLFSLSRELKLLDLSRN 180
QY 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNHLRSVAPGAFDR 240
DB 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNHLRSVAPGAFDR 240
QY 241 LGNLSLTLGSLNLSLPPALFLHVSVSRITLFPENPLELDPVLFCEMAGLRELMNGT 300
DB 241 LGNLSLTLGSLNLSLPPALFLHVSVSRITLFPENPLELDPVLFCEMAGLRELMNGT 300
QY 301 HLSTLPAAAFNLSGLQTLGTRNPRLSALPRGVFQGLRELRVLAHTNALAEURDDALR 360
DB 301 HLSTLPAAAFNLSGLQTLGTRNPRLSALPRGVFQGLRELRVLAHTNALAEURDDALR 360
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389
DB 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 2
US-10-037-417-105
; Sequence 105, Application US/10037417
; Publication No. US20040052806A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Sheno, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eison, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/10/037,417
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 105
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-037-417-105
Query Match 100.0%; Score 1938; DB 15; Length 567;
Best Local Similarity 100.0%; Pred. No. 4.6e-156; Indels 0; Gaps 0;
Matches 389; Conservative 0; Mismatches 0;
QY 1 MLRSALLSAVLALLRAQFPCKTKCVVRDAAQCSGGSVAHIAELGLPTNLTHILFRM 60
DB 1 MLRSALLSAVLALLRAQFPCKTKCVVRDAAQCSGGSVAHIAELGLPTNLTHILFRM 60
QY 61 DQGI LRNHSFGMTVLQRLMLSDSHISAIDPCTNDLVKLTLETRNKISRLPAILDK 120
DB 61 DQGI LRNHSFGMTVLQRLMLSDSHISAIDPCTNDLVKLTLETRNKISRLPAILDK 120
QY 121 MVLLEQLFDHNAFLDQNLFOQLRNQLGELGNQNLQSLFPLANLFSLSRELKLLDLSRN 180
DB 121 MVLLEQLFDHNAFLDQNLFOQLRNQLGELGNQNLQSLFPLANLFSLSRELKLLDLSRN 180
QY 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNHLRSVAPGAFDR 240
DB 181 NLTHLPKGLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRNHLRSVAPGAFDR 240
QY 241 LGNLSLTLGSLNLSLPPALFLHVSVSRITLFPENPLELDPVLFCEMAGLRELMNGT 300
DB 241 LGNLSLTLGSLNLSLPPALFLHVSVSRITLFPENPLELDPVLFCEMAGLRELMNGT 300
QY 301 HLSTLPAAAFNLSGLQTLGTRNPRLSALPRGVFQGLRELRVLAHTNALAEURDDALR 360
DB 301 HLSTLPAAAFNLSGLQTLGTRNPRLSALPRGVFQGLRELRVLAHTNALAEURDDALR 360
QY 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389
DB 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 3
US-10-037-417-106
; Sequence 106, Application US/10037417
; Publication No. US20040052806A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Sheno, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E

```
; APPLICANT: Eisen, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/10/037,417
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-037-417-106

Query Match          99.2%; Score 1923; DB 15; Length 567;
Best Local Similarity 99.2%; Pred. No. 8,78-155; Indels 0; Gaps 0;
Matches 386; Conservative 0; Mismatches 3;

Qy 1 MLRSALLSAVLALLRAQFPCKTKCKVVRDAAQCGSGVAHIAELGLPTNLTHILFRM 60
Db 1 MLRSALLSAVLALLRAQFPCKTKCKVVRDAAQCGSGVAHIAELGLPTNLTHILFRM 60
Qy 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTENDLVKLTKLRLTKNKSRLPRAILDK 120
Db 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTENDLVKLTKLRLTKNKSRLPRAILDK 120
Qy 121 MVLEQLFLDHNALRDLDQNLFOQLRNQLQELGNQNLSPFLPANLFSLSRELKLLDLSRN 180
Db 121 MVLEQLFLDHNALRDLDQNLFOQLRNQLQELGNQNLSPFLPANLFSLSRELKLLDLSRN 180
Qy 181 NLTHLPKLLGAQVKLEKLLYSNQLTSVDSGLLSNLCALTELELERNHLRSVAPGAFDR 240
Db 181 NLTHLPKLLGAQVKLEKLLYSNQLTSVDSGLLSNLCALTELELERNHLRSVAPGAFDR 240
Qy 241 LGNLSLTLGSLNLSLPPALFLHVSVSRILTLFENPLELDPVLFGEAGLRELWNGT 300
Db 241 LGNLSLTLGSLNLSLPPALFLHVSVSRILTLFENPLELDPVLFGEAGLRELWNGT 300
Qy 301 HLSTLPAAARNLSGLTGLTRNPRLSALPRGVFQGLRELRVLAHTNALAEIRDALR 360
Db 301 HLSTLPAAARNLSGLTGLTRNPRLSALPRGVFQGLRELRVLAHTNALAEIRDALR 360
Qy 361 GLGHLRQVSLRHNRRLRALPRTLPRNLSSL 389
Db 361 GLGHLRQVSLRHNRRLRALPRTLPRNLSSL 389

RESULT 4
US-10-080-334-228
; Sequence 228, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G

; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spyttek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Beha A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 228
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-080-334-228

Query Match          90.2%; Score 1749; DB 15; Length 567;
Best Local Similarity 89.2%; Pred. No. 5,58-140; Indels 0; Gaps 0;
Matches 347; Conservative 24; Mismatches 18;

Qy 1 MLRSALLSAVLALLRAQFPCKTKCKVVRDAAQCGSGVAHIAELGLPTNLTHILFRM 60
Db 1 MLRSALLSAVLALLRAQFPCKTKCKVVRDAAQCGSGVAHIAELGLPTNLTHILFRM 60
Qy 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTENDLVKLTKLRLTKNKSRLPRAILDK 120
Db 61 DQGLRNHPSFGMTVLQRLMLSDSHISAIDPGTENDLVKLTKLRLTKNKSRLPRAILDK 120
```

Db 61 DRGVLSQHSFSGMTVQLQRLMLSDSHISAIDPGTFNDLVKLTIRLRNKHSHLPRAILDK 120
Qy 121 MVLLEQLFDHNAERLDQNLFOQLRNLOELGNQQLSFLPANLFFSSIRELKLDLSRN 180
Db 121 MVLLEQLFDHNAERLDQNLFOQLRNLOELGNQQLSFLPANLFFSSIRELKLDLSRN 180
Qy 181 NLTHLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 240
Db 181 NLTHLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 240
Qy 241 LGNLSLTLGNLLESPPALFLHVSVSRLLTTFENPLEPDLVFGEMAGRELWLNGT 300
Db 241 LGNLSLTLGNLLESPPALFLHVSVSRLLTTFENPLEPDLVFGEMAGRELWLNGT 300
Qy 301 HLTLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 360
Db 301 HLTLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 360
Qy 361 GLGHLRQVSLRHNRLRALPRLFRNLSSL 389
Db 361 GLGHLRQVSLRHNRLRALPRLFRNLSSL 389

RESULT 5

US-09-782-980-61
; Sequence 61, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: MacBeath, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: STIMULANT PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: STIMULANT PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 09/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/672,721
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/049,799
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61

; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-980-61
Query Match 76.2%; Score 1477; DB 9; Length 560;
Best Local Similarity 76.1%; Pred. No. 7,4e-117;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;
Qy 1 MRSALLSAVLALLRAQPPPCPKCKCVVRDAAQCGSVAHIAELGLPTNLTHILLFRM 60
Db 1 MVRGTLCAVLGLLRAQPPPCPKCKCVVRDAAQCGSVAHIAELGLPTNLTHILLFRM 60
Qy 61 DOGILRNHFSFGMTVQLQRLMLSDSHISAIDPGTFNDLVKLTIRLRNKHSHLPRAILDK 120
Db 61 GRGVLSQHSFSGMTVQLQRLMLSDSHISAIDPGTFNDLVKLTIRLRNKHSHLPRAILDK 120
Qy 121 MVLLEQLFDHNAERLDQNLFOQLRNLOELGNQQLSFLPANLFFSSIRELKLDLSRN 180
Db 121 MVLLEQLFDHNAERLDQNLFOQLRNLOELGNQQLSFLPANLFFSSIRELKLDLSRN 180
Qy 181 NLTHLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 240
Db 181 NLTHLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 240
Qy 241 LGNLSLTLGNLLESPPALFLHVSVSRLLTTFENPLEPDLVFGEMAGRELWLNGT 300
Db 241 LGNLSLTLGNLLESPPALFLHVSVSRLLTTFENPLEPDLVFGEMAGRELWLNGT 300
Qy 301 HLTLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 360
Db 301 HLTLPKGLLGAQVKLEKLLYSNOLTSVDSGLSNLGLTALTELRLRNHLRSVAPGAFDR 360
Qy 361 GLGHLRQVSLRHNRLRALPRLFRNLSSL 389
Db 361 GLGHLRQVSLRHNRLRALPRLFRNLSSL 389

RESULT 6

US-09-775-803-14
; Sequence 14, Application US/09775803
; Publication No. US20030167487A1
; GENERAL INFORMATION:
; APPLICANT: COR Therapeutics, Inc.
; APPLICANT: Ramakrishnan, Vanitha
; APPLICANT: Phillips, David
; TITLE OF INVENTION: Transgenic Animals Having a Modified Glycoprotein V
; TITLE OF INVENTION: Gene and Methods for Their Use
; FILE REFERENCE: 44481-5044-US
; CURRENT APPLICATION NUMBER: US/09/775,803
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: PCT/US99/17594
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: US 60/109,797
; PRIOR FILING DATE: 1998-08-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-775-803-14
Query Match 76.2%; Score 1477; DB 10; Length 560;
Best Local Similarity 76.1%; Pred. No. 7,4e-117;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;
Qy 1 MRSALLSAVLALLRAQPPPCPKCKCVVRDAAQCGSVAHIAELGLPTNLTHILLFRM 60
Db 1 MVRGTLCAVLGLLRAQPPPCPKCKCVVRDAAQCGSVAHIAELGLPTNLTHILLFRM 60
Qy 61 DOGILRNHFSFGMTVQLQRLMLSDSHISAIDPGTFNDLVKLTIRLRNKHSHLPRAILDK 120
Db 61 GRGVLSQHSFSGMTVQLQRLMLSDSHISAIDPGTFNDLVKLTIRLRNKHSHLPRAILDK 120

Db 61 GRGVLSQSFSGMTVQLRLMTSDSHISAVAPGTSDIUKTKTLELSRNKITHLPALLDK 120
QY 121 MVLLQFLDHNALRDQNLFOQLRNLOELGNQNLSPALPNSLSRELKLLDLSRN 180
Db 121 MVLLQFLDHNALRGIDQNMFKLVNLOELALNQQLDFLPASLFTNLENKLLDLSGN 180
QY 181 NTHLPKGLGCAQVKLEKLLYSNQLTSVDSGLLSNGLTALTELERNHLSRVAPGAFDR 240
Db 181 NTHLPKGLGCAQVKLEKLLYSNQLTSVDSGLLSNGLTALTELERNHLSRVAPGAFDR 240
QY 241 LGNLSLTLSSLTSLNLSLPPALFLHVSVSRITLTFENPLELDPVLFGEMAGLRELWNGT 300
Db 241 LPNLSLTLSSLTSLNLSLPPALFLHVSVSRITLTFENPLELDPVLFGEMAGLRELWNGT 300
QY 301 HLSTLPAAAFNLGLQTLGTRNPRLSALPRGVFGQIRLVRVLAHTNALAEIRDDALR 360
Db 301 QLRTLPAAAFNLGLQTLGTRNPRLSALPRGVFGQIRLVRVLAHTNALAEIRDDALR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
Db 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389

RESULT 7

US-10-212-499-2
; Sequence 2, Application US/10212499
; Publication No. US20030135036A1

GENERAL INFORMATION:

APPLICANT: Lanza, Francois
Phillips, David R.
Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan Lewis & Bockius LLP
STREET: 1800 M St., NW
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/212,499
FILING DATE: 06-Aug-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/560,814
FILING DATE: 2000-04-28
APPLICATION NUMBER: US/08/089,455
FILING DATE: 1993-07-09
APPLICATION NUMBER: US/08/195,006
FILING DATE: 1994-02-10
APPLICATION NUMBER: US/08/884,571
FILING DATE: 1997-08-27

ATTORNEY/AGENT INFORMATION:

NAME: Reid G. Adler
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 44481-5018-04-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 560 amino
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-212-499-2

Query Match

76.2%; Score 1477; DB 14; Length 560;

Best Local Similarity 76.1%; Pred. No. 7.4e-117;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MERSALTSALVALLRAQPPCKTKCVVRDAAQCGSGVAHIAELGLPTNLTHTLLPRM 60
Db 1 MURGTLLCAGVGLLRAQFFPCPPACKCVFRDAAQCGGVDARISALGPTNLTHTLLPRM 60
QY 61 DQGLRNHSFGMTVQLRLMTSDSHISAVAPGTSDIUKTKTLELSRNKITHLPALLDK 120
Db 61 GRGVLSQSFSGMTVQLRLMTSDSHISAVAPGTSDIUKTKTLELSRNKITHLPALLDK 120
QY 121 MVLLQFLDHNALRDQNLFOQLRNLOELGNQNLSPALPNSLSRELKLLDLSRN 180
Db 121 MVLLQFLDHNALRGIDQNMFKLVNLOELALNQQLDFLPASLFTNLENKLLDLSGN 180
QY 181 NTHLPKGLGCAQVKLEKLLYSNQLTSVDSGLLSNGLTALTELERNHLSRVAPGAFDR 240
Db 181 NTHLPKGLGCAQVKLEKLLYSNQLTSVDSGLLSNGLTALTELERNHLSRVAPGAFDR 240
QY 241 LGNLSLTLSSLTSLNLSLPPALFLHVSVSRITLTFENPLELDPVLFGEMAGLRELWNGT 300
Db 241 LPNLSLTLSSLTSLNLSLPPALFLHVSVSRITLTFENPLELDPVLFGEMAGLRELWNGT 300
QY 301 HLSTLPAAAFNLGLQTLGTRNPRLSALPRGVFGQIRLVRVLAHTNALAEIRDDALR 360
Db 301 QLRTLPAAAFNLGLQTLGTRNPRLSALPRGVFGQIRLVRVLAHTNALAEIRDDALR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
Db 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389

RESULT 8

US-10-806-018-61

; Sequence 61, Application US/10806018
; Publication No. US20040176296A1

GENERAL INFORMATION:

APPLICANT: Khodadoust, Mehran M.
APPLICANT: MacBeth, Kyle J.
APPLICANT: Busfield, Samantha J.
APPLICANT: McCarthy, Sean A.
APPLICANT: Holtzman, Douglas A.
APPLICANT: Gu, Wei
APPLICANT: White, David
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIPE, TRASH, BDSF, LRSG, AND
TITLE OF INVENTION: STMT PROTEIN AND NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: MNI-121CP
CURRENT APPLICATION NUMBER: US/10/806,018
CURRENT FILING DATE: 2004-03-22
PRIOR APPLICATION NUMBER: US/09/782,980
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: PCT/US00/02125
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 09/448,076
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: 09/276,400
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: 60/117,580
PRIOR FILING DATE: 1999-01-27
PRIOR APPLICATION NUMBER: 09/014,195
PRIOR FILING DATE: 1998-01-27
PRIOR APPLICATION NUMBER: 09/014,348
PRIOR FILING DATE: 1998-01-27
PRIOR APPLICATION NUMBER: 09/086,892
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 09/296,208
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 09/063,950
PRIOR FILING DATE: 1998-04-21

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 176

SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 61
 LENGTH: 560
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-806-018-61

Query Match 76.2%; Score 1477; DB 16; Length 560;
 Best Local Similarity 76.1%; Pred. No. 7, 4e-117;
 Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MRSALLSAVALLRAQPPCPKTCVVRDAAQCGGSAHIAELGPTNLTTHILPRM 60
 DB 1 MLRGTLCAVLGLRAQPPCPACVCFRDAACQCGGSAVARISALGLPTNLTTHILFGM 60

QY 61 DOGLRNHSGMTVLQRLMLSDSHISADPGTFNDLVKLTRLTRNKISRLPAIDK 120
 DB 61 GRGVLOSQSGMTVLQRLMISDHSISAVAPGTFSDLIKLTLSRNKITHLPGALLDK 120

QY 121 MVLEQLFDHNAALDQNLFOQLNQLGELGNQNSFLPANLFSRLREIKLKLDSRN 180
 DB 121 MVLEQLFDHNAALDQNLFOQLNQLGELGNQNSFLPANLFSRLREIKLKLDSGN 180

QY 181 NLTHLPKGLGAOVKLEKLLYSNOLTSVDSGLNSLGNLGTALTELRLRNHLSVAPGAFDR 240
 DB 181 NLTHLPKGLGAOVKLEKLLYSNOLTSVDSGLNSLGNLGTALTELRLRNHLSVAPGAFDR 240

QY 241 LGNLSLTLGSLNLESPPALFHVSVSRLLTFENPLEELPDVLFGEAGRELWNGT 300
 DB 241 LPNLSLTLGSLNLESPPALFHVSVSRLLTFENPLEELPDVLFGEAGRELWNGT 300

QY 301 HLSTLPAAPRNLSGLTGLTENPRLSALPRGVFGELRELRLVLAHTNALAELEDDALR 360
 DB 301 QLRTLPAAPRNLSRRLYGLVTLSPLRSALPGAFQGLGELQVLAHNSGLTALPDGLLR 360

QY 361 GLGHLROVSLRNRLRALPRTLPRNLSSL 389
 DB 361 GLGHLROVSLRNRLRALPRTLPRNLSSL 389

RESULT 9

US-10-312-311-21
 ; Sequence 21, Application US/10312311
 ; Publication No. US20040072173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Boyle, Bryan
 ; APPLICANT: Yeung, George
 ; APPLICANT: Mize, Nancy
 ; APPLICANT: Aterburn, Matthew
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Wang, Meng-Yun
 ; APPLICANT: Chen, Lichuan
 ; APPLICANT: Yang, Yea-Huey
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO LEUCINE-RICH REPEAT
 ; TITLE OF INVENTION: PROTEIN-LIKE (LRR
 ; FILE REFERENCE: HYS-26/ 21272-032
 ; CURRENT APPLICATION NUMBER: US/10/312,311
 ; CURRENT FILING DATE: 2002-12-19
 ; PRIOR APPLICATION NUMBER: 09/672,221
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 21
 ; LENGTH: 368
 ; TYPE: PRT

ORGANISM: Homo sapiens
 US-10-312-311-21

Query Match 62.0%; Score 1201; DB 15; Length 368;
 Best Local Similarity 76.4%; Pred. No. 1.3e-93;
 Matches 243; Conservative 32; Mismatches 43; Indels 0; Gaps 0;

QY 72 GMTVLQRLMLSDSHISADPGTFNDLVKLTRLTRNKISRLPAIDKMLLEQLFLDH 131
 DB 1 GMTVLQRLMISDHSISAVAPGTFSDLIKLTLSRNKITHLPGALLDKMLLEQLFLDH 60

QY 132 NALRDLQNLFOQLNQLGELGNQNSFLPANLFSRLREIKLKLDSRNLTTHLPKGLLG 191
 DB 61 NALRGIDQNMFOQLNQLGELGNQNSFLPANLFSRLREIKLKLDSRNLTTHLPKGLLG 120

QY 192 AOVKLEKLLYSNOLTSVDSGLNSLGNLGTALTELRLRNHLSVAPGAFDRNLSSLTSG 251
 DB 121 AAKERLRLHNSRLVSDSGLNSLGNLGTALTELRLRNHLSVAPGAFDRNLSSLTSG 190

QY 252 NLESPLPALFHVSVSRLLTFENPLEELPDVLFGEAGRELWNGTTHLSTLPAAPR 311
 DB 181 NLAFLPSALFHVSHNLTLLTFENPLAELPGVLFGEAGRELWNGTTHLSTLPAAPR 240

QY 312 NLSGLOTGLTENPRLSALPRGVFGELRELRLVLAHTNALAELEDDALRGHLROVSLR 371
 DB 241 NLSRLYGLVTLSPLRSALPGAFQGLGELQVLAHNSGLTALPDGLLRGLKROVSLR 300

QY 372 HNRRLALPRTLPRNLSSL 389
 DB 301 RNRRLALPRTLPRNLSSL 318

RESULT 10

US-10-295-027-410
 ; Sequence 410, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Bos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.

APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
TITLE OF INVENTION: Using the Same
FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/276,677
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23

US-10-080-334-76

Query Match 31.3%; Score 607; DB 15; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTCVVYDAAQCSGSGVAHIAELGLPTNLTHILLFMDQGIILRNHFSFGMTVLQRLM 80
DB 25 CSECTCSASQVECTGARIVAV-PTPLPWANMSQILNTHITELNESFFLNISALIALR 83
QY 81 LSDSHISAIDPGTFNDVLKTLRLTRNKISLPRALDKMVLLEQLFLDHNALRDLQDN 140
DB 84 IEKNVLSRTPGAFNLGSLRYLSLANNKLQVLPICLFGQLDSLSLSLLSNQLLQIOPA 143
QY 141 LPQQLNLQELGNQQLSFLPANLFSLSRELKLDLSNNLTHLPKGLGAGVLEKLL 200
DB 144 HFSQCNSLQELQHNHLSYIDGADHVLVGTUKNLGKNSLTHISPRVFOHGLNQLVR 203
QY 201 LYENQLTSDVSGLLSGLNGLALTELRLRNHLRSVAPGAFRLGNLSLTLSGNLLSLLPPA 260
DB 204 LYENRLTDIPMGFFDGLVNLQELALQONQIGLLSPGLFNNHNLQRLVLSNNHISQLPPS 263
QY 261 LFHVSVSRLLTFENPLEPLPVLGEMAGIRELMNGLNTHLSTLPAAFRNLGSLQTLG 320
DB 264 IFNQLPQLNRLTLFGNSKLSLGIFGPMFNRELMLYDNHLSUPDNVFSNLRQLQVLI 323

US-10-080-334-225

RESULT 13
US-10-080-334-225
Sequence 225, Application US/10080334
Publication No. US20040002584A1
GENERAL INFORMATION:
APPLICANT: Pena, Carol E. A.
APPLICANT: Shimkets, Richard A.
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A.
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Guo, Xiaojia
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna
APPLICANT: Tchernev, Velizar T
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A
APPLICANT: Padigaru, Muralidhara
APPLICANT: Liu, Xiaohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
TITLE OF INVENTION: Using the Same
FILE REFERENCE: 21402-275
CURRENT APPLICATION NUMBER: US/10/080,334
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/330,307
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/278,796
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 60/281,521
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/276,677
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/311,595
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/270,220
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/274,295
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/318,526
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/286,548
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/291,765
PRIOR FILING DATE: 2001-05-17
PRIOR APPLICATION NUMBER: 60/270,797
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/276,400
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/270,810
PRIOR FILING DATE: 2001-02-23

```
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 225
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-225

Query Match      31.3%; Score 607; DB 15; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAACQSGGSAHIAELGLPTNLTHILLFRMDQGLRNHHSFGMTVLQRLM 80
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPFWNAMSQILNTHITELNESFPLNISALIALR 83
QY 81 LSDSHISAIDPGTFNDLVKLTIRLNKISRLPRAILDKMWLLQQLFDHNAERDLDQ 140
Db 84 IERNELSRITPGAFNRLGSLRYLSLANNKQLVLPGLFQGLDSLESLLSSNQLLOQPA 143
QY 141 LFOQLRNLOELGNQNSFLPANLFSLSRELKLDLSNNLTHLPKGLGGAQVKLEKL 200
Db 144 HFSQCSNLKELQHGHNLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFOHGLNQLVR 203
QY 201 LYSNQLTSVDSGLLSNLGALTELRLRNHLRSVAPGAFDRGLNLSLTLISGNLLESPPA 260
Db 204 LYENELTDIPMGTFDGLVNLQELALQQOIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263
QY 261 LFLHVSVSRLTLPENPLEELPDVLFEGMAGRLWLNGLTHLSTLPAAAFNLSGLQTLG 320
Db 264 IFMQLPQLNRLLTFGNSLKLSLGLFGFPMNRLWLYDNHISLSDPNVFSNLRQLQVLI 323
QY 321 LTRNPRLSALPRGVFGQLRELRLVLAHTNALAEALRDALRGHLRQVSLRHNRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLGDNVFRMLANLQNSLQNNRLRLQPLG 382
QY 381 TLFRLNSSL 389
Db 383 NIFANVNGL 391

RESULT 15
US-10-080-334-224
; Sequence 224, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Kerkuda, Ramesh
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980

; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 225
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-225

Query Match      31.3%; Score 607; DB 15; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAACQSGGSAHIAELGLPTNLTHILLFRMDQGLRNHHSFGMTVLQRLM 80
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPFWNAMSQILNTHITELNESFPLNISALIALR 83
QY 81 LSDSHISAIDPGTFNDLVKLTIRLNKISRLPRAILDKMWLLQQLFDHNAERDLDQ 140
Db 84 IERNELSRITPGAFNRLGSLRYLSLANNKQLVLPGLFQGLDSLESLLSSNQLLOQPA 143
QY 141 LFOQLRNLOELGNQNSFLPANLFSLSRELKLDLSNNLTHLPKGLGGAQVKLEKL 200
Db 144 HFSQCSNLKELQHGHNLEYIPDGAFDHLVGLTKNLGKNSLTHISPRVFOHGLNQLVR 203
QY 201 LYSNQLTSVDSGLLSNLGALTELRLRNHLRSVAPGAFDRGLNLSLTLISGNLLESPPA 260
Db 204 LYENELTDIPMGTFDGLVNLQELALQQOIGLLSPGLFHNHNLQRLYLSNNHISQLPPS 263
QY 261 LFLHVSVSRLTLPENPLEELPDVLFEGMAGRLWLNGLTHLSTLPAAAFNLSGLQTLG 320
Db 264 IFMQLPQLNRLLTFGNSLKLSLGLFGFPMNRLWLYDNHISLSDPNVFSNLRQLQVLI 323
QY 321 LTRNPRLSALPRGVFGQLRELRLVLAHTNALAEALRDALRGHLRQVSLRHNRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLGDNVFRMLANLQNSLQNNRLRLQPLG 382
QY 381 TLFRLNSSL 389
Db 383 NIFANVNGL 391

RESULT 14
US-10-258-951-43
; Sequence 43, Application US/10258951
; Publication No. US2004003504A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Xie, Qing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GFS0025
; CURRENT APPLICATION NUMBER: US/10/258,951
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: PCT/US01/13360
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/199,963
; PRIOR FILING DATE: 2000-04-27/203,336
; PRIOR APPLICATION NUMBER: 60/207,087
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/207,546
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 581
; TYPE: PRT
```

```
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 224
; LENGTH: 787
; TYPE: CDS
; ORGANISM: Homo sapiens
US-10-080-334-224

Query Match      31.3%; Score 607; DB 15; Length 787;
Best Local Similarity 37.7%; Pred. No. 1.2e-42;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTKCVVRDAQCQSGSVAHIAELGPTNLTHILFRMDQGIILRNHFSGTMVLQRLM 80
DB 231 CPSECTCSRASQVECTGARIVAV-PTPLPWAMSLQILNTHITELNESPFNLISALIALR 289

QY 81 LSDSHISAIDPGTFNDLVKLTIRLTKRNKISRLPRAILDPMVLLBQLFDHNAIRDLDQN 140
DB 290 IEKNLSRITPGAPRNLSRLVSLANNKLQVLPGLFQGLDLSLESLSSNQLIQIPA 349

QY 141 LFCQRLNQLGILNQNLSELPANLFSLSRELKLLDLSNNLTHLPKGLGAQVLEKLL 200
DB 350 HFSQCSNKLQHLGNHLEYIPGADPHLVGLTKLNGKNSLTHISPRVQHLGNLQVLR 409

QY 201 LYSNQLTSVDSGLSNLGALETLRLRNHLRSVAPGAFDRGLNLSLTLGSLNLSLPPA 260
DB 410 LYENRLTDIPMGTFDGLVNLQELALQONQIGLISPLPHNNLQRLVLSNNHISQLPPS 469

QY 261 LFLHVSSVSLTLPENLESLPDVLGEMAGIRELMNGTHLSTLPAAFRNLISGLQTLG 320
DB 470 IFMQLPQNLRLTLFGNSKELSLGIFGPMNLRKELWLYONHLYSSLPDNDVFSNLRQLVLI 529

QY 321 LTRNPRLSALPRGVQFQGLRELRLVLAHTNALAEIRDALRGHLRQVSLRNRLRALPR 380
DB 530 LSRN-QISFTSPGAFNGLTELRSLHTNALQDLGDNVFRMLANLQNISLQNNRLRQLPG 588

QY 381 TLFFNLSSL 389
DB 589 NIFANVNGL 597
```

Search completed: November 29, 2004, 13:18:03
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 12:59:28 ; Search time 109 Seconds
(without alignments)
2053.401 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389

Perfect score: 1938

Sequence: 1 MLRSALLSAVLALLRAQPEP.....LRHNRLEALPRTLFRLNSSL 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1938	100.0	567	2 Q9QZU3	Q9qzu3 mus musculus
2	1923	99.2	567	1 GPV_MOUSE	O08742 mus musculus
3	1749	90.2	567	1 GPV_RAT	O08770 rattus norv
4	1477	76.2	560	1 GPV_HUMAN	P40197 homo sapien
5	782	40.4	637	2 Q6DCV7	O6dcv7 xenopus lae
6	607	31.3	581	1 LR15_HUMAN	O8tf66 homo sapien
7	607	31.3	581	2 Q7RNF7	Q7rtn7 homo sapien
8	600	31.0	579	2 Q80X72	Q80x72 mus musculus
9	583	30.1	578	1 LR15_RAT	Q8f5m3 rattus norv
10	541.5	27.9	391	2 Q9D3K0	Q9d3k0 m mus muscu
11	503.5	26.0	562	2 Q8R113	O8r113 mus musculus
12	503.5	26.0	570	2 Q9DBB9	O9dbb9 mus musculus
13	445	23.0	557	2 Q8N5V4	O8n5v4 homo sapien
14	445	23.0	560	2 Q86S04	Q86s04 homo sapien
15	419.5	21.6	605	1 ALS_HUMAN	P35858 homo sapien
16	416.5	21.5	536	1 CBP8_HUMAN	P22792 homo sapien
17	416.5	21.3	605	1 Q8TAY0	O8tay0 homo sapien
18	413.5	21.3	605	1 ALS_PAPHA	O2833 papio hamad
19	397	20.5	3127	2 Q7PTD4	Q7ptd4 anopheles g
20	384	19.8	1000	2 Q7Q888	Q7qf88 anopheles g
21	382.5	19.7	603	1 ALS_RAT	P35859 rattus norv
22	382	19.7	603	1 ALS_MOUSE	F70389 mus musculus
23	382	19.7	603	2 AAF69482	Aaf69482 mus muscu
24	382	19.7	687	2 Q9J1L0	Q9j1l0 mus musculus
25	381.5	19.7	603	2 Q70211	O70211 rattus norv
26	379.5	19.6	1531	1 SLT1_RAT	O8279 rattus norv
27	379.5	19.6	1534	1 SLT1_HUMAN	O75093 homo sapien
28	379	19.6	417	2 Q6EAJ7	Q6eaj7 petromyzon
29	375.5	19.4	1531	1 SLT1_MOUSE	O80tr4 mus musculus
30	372.5	19.2	1058	2 Q76FN7	Q76fn7 tachypneus
31	372.5	19.2	1058	2 BAD12073	Bad12073 tachypneu

32	370	19.1	890	2 Q7Q941	Q7q941 anopheles g
33	369.5	19.1	766	1 SLT2_RAT	Q9wvc1 rattus norv
34	367.5	19.0	907	1 LGR5_MOUSE	Q9zlp4 mus musculus
35	367.5	19.0	1521	1 SLT2_MOUSE	Q9rlb9 mus musculus
36	366.5	18.9	1392	2 Q9VAD1	Q9vad1 drosophila
37	366	18.9	1523	1 SLT3_RAT	O88280 rattus norv
38	365	18.8	451	2 Q7Q1S1	Q7q1s1 anopheles g
39	364.5	18.8	615	2 Q9VZ84	Q9vz84 drosophila
40	364.5	18.8	1529	2 Q7ZX12	Q7zx12 xenopus lae
41	362	18.7	1523	1 SLT3_MOUSE	Q9wvba4 mus musculus
42	361.5	18.7	1529	1 SLT2_HUMAN	O94813 homo sapien
43	361.5	18.7	1530	2 Q90WZ3	Q90wz3 xenopus lae
44	361	18.6	1173	2 Q7QHH1	Q7qhh1 anopheles g
45	360	18.6	737	2 Q96SM3	Q96sm3 caenorhabdi

ALIGNMENTS

RESULT 1

Q9QZU3 PRELIMINARY; PRT; 567 AA.

ID Q9QZU3 AC Q9QZU3 DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Platelet glycoprotein V (Fragment) (Mus musculus 13 days embryo male
DE testis cDNA, RIKEN full-length enriched library, clone:6030400G03
DE Product:glycoprotein 5 (platelet), full insert sequence)
DE (Fragment)
GN Name=Gp5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=20027548; PubMed=10557321;
RA Ramakrishnan V., Reeves P.S., DeGuzman F., Deshpande U.,
RA Ministri-Madrid K., DuBridge R.B., Phillips D.R.;
RT "Increased thrombin responsiveness in platelets from mice lacking
RT glycoprotein V.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13336-13341(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE-Testis;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE-Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE-Testis;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE-Testis;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [6]

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuho H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
 RA Okazaki Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Yoneda Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [7]

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume M.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Horii F., Imozaki K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki R., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tonaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF163101; AA08787.1; -
 DR EMBL; AK031272; BAC27331.1; -
 DR HSP; Q9BZ86; 102N.

DR MGD; MGI:1096363; Gp5.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR_14.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PRO00019; LEURICHRPT.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00369; LRR_Typ; 10.
 FT NON TER 567
 SQ SEQUENCE 567 AA; 63381 MW; E6AP470855ACB5C9 CRC64;

Query Match 100.0%; Score 1938; DB 2; Length 567;
 Best Local Similarity 100.0%; Pred. No. 2.3e-122;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSALLSAVIALRAQFPFCPTKCKVVRDAAQCSCGGSVAHIAELGLPTNLTHILLFRM 60
 DB 1 MRSALLSAVIALRAQFPFCPTKCKVVRDAAQCSCGGSVAHIAELGLPTNLTHILLFRM 60
 QY 61 DQGLRNHSGGTVLQRLMDSHISADPGFNDVLKTLRTNKISLPAILDK 120
 DB 61 DQGLRNHSGGTVLQRLMDSHISADPGFNDVLKTLRTNKISLPAILDK 120
 QY 121 MVLLLEQLFDHNRDLQDLFQRLNLQELGNLQNLQSLFANLFSSRLKLLDLGRN 180
 DB 121 MVLLLEQLFDHNRDLQDLFQRLNLQELGNLQNLQSLFANLFSSRLKLLDLGRN 180
 QY 181 NLTHLPKGLGAQVKELLYNQTSVDSGLSNLGLTTLRLERHLSRVAQADR 240
 DB 181 NLTHLPKGLGAQVKELLYNQTSVDSGLSNLGLTTLRLERHLSRVAQADR 240
 QY 241 LGNLSSTLSGNLLESPPALFLHVSVSRLTLFENPLEELPDVLFGEAGRLWLNGT 300
 DB 241 LGNLSSTLSGNLLESPPALFLHVSVSRLTLFENPLEELPDVLFGEAGRLWLNGT 300
 QY 301 HSLTLPAAAFNLSGLTGLTNPRISALPRGVFGRLRLVLAHTNALAELEDDALR 360
 DB 301 HSLTLPAAAFNLSGLTGLTNPRISALPRGVFGRLRLVLAHTNALAELEDDALR 360

QY 361 GLCHLRQVSLRHRNLRALPRTLFRNLSSL 389
 DB 361 GLCHLRQVSLRHRNLRALPRTLFRNLSSL 389

RESULT 2

GPV_MOUSE STANDARD; PRT; 567 AA.
 AC 008742;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Platelet glycoprotein V precursor (GPV) (CD42D).
 GN Names-Gp5;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=97275336; PubMed=9129030;
 RA Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S., Grunert P.,
 RA Loew D., van Dorsselaer A., Cazenave J.-P., Lanza P.,
 RT "Gene cloning of rat and mouse platelet glycoprotein V: identification
 RT of megakaryocyte-specific promoters and demonstration of functional
 RT thrombin cleavage."
 RL Blood 89:3253-3262(1997).
 CC -1- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand
 CC factor receptor and mediates von Willebrand factor-dependent
 CC platelet adhesion to blood vessels. The adhesion of platelets to
 CC injured vascular surfaces in the arterial circulation is a
 CC critical initiating event in hemostasis (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; Z69595; CAA93441.1; -
 DR HSP; Q9BZ86; 102N.
 DR MGD; MGI:1096363; Gp5.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR Pfam; PF00560; LRR; 13.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PRO00019; LEURICHRPT.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00369; LRR_Typ; 10.
 KW Blood coagulation; Cell adhesion; Glycoprotein; Leucine-rich repeat;
 KW Platelet; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 16
 FT Potential.
 FT CHAIN 17 567
 FT DOMAIN 17 522
 FT TRANSMEM 523 543
 FT DOMAIN 544 567
 FT REPEAT 73 96
 FT REPEAT 97 120
 FT REPEAT 122 144
 FT REPEAT 145 168
 FT REPEAT 170 192
 FT REPEAT 194 216
 FT REPEAT 217 240
 FT REPEAT 241 264
 FT REPEAT 265 288
 FT REPEAT 289 312
 FT LRR 1.
 FT LRR 2.
 FT LRR 3.
 FT LRR 4.
 FT LRR 5.
 FT LRR 6.
 FT LRR 7.
 FT LRR 8.
 FT LRR 9.
 FT LRR 10.


```

FT REPEAT 314 337 LRR 11.
FT REPEAT 338 361 LRR 12.
FT REPEAT 362 385 LRR 13.
FT REPEAT 386 409 LRR 14.
FT CARBOHYD 51 51 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 67 67 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 181 181 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 243 243 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 312 312 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 385 385 N-linked (GlcNAc. . .) (Potential).
SQ SEQUENCE 567 AA; 63467 MW; C48643AA73967AD7D CRC64;

Query Match 99.2%; Score 1923; DB 1; Length 567;
Best Local Similarity 99.2%; Pred. No. 2.4e-121;
Matches 386; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPKCKCVRDAAQCSCGSAHAIAELGLPTNLTHILLFRM 60
DB 1 MLRSALLSAVLPLLRQAPFPCKCKCVRDAAQCSCGSAHAIAELGLPTNLTHILLFRM 60
QY 61 DOGILRNHSPSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTKNLSRLPRAILDK 120
DB 61 DOGILRNHSPSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTKNLSRLPRAILDK 120
QY 121 MVLEQFLDHNALRDLDQNLFOQLRNQLGELGNQNLQSLFPLNLFSSRELKLLDLSRN 180
DB 121 MVLEQFLDHNALRDLDQNLFOQLRNQLGELGNQNLQSLFPLNLFSSRELKLLDLSRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNIGALTEURLERNHLRSVAPGAFDR 240
DB 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNIGALTEURLERNHLRSVAPGAFDR 240
QY 241 LGLNLSLTLGNNLESPPALFLHVSSVSLTTFENPLELPDVLFGEMAGRELWNGT 300
DB 241 LGLNLSLTLGNNLESPPALFLHVSSVSLTTFENPLELPDVLFGEMAGRELWNGT 300
QY 301 HLSTLPAARFNSGLQTLGTRNPLRSALPRGVQGLRELRLVGLHTNALAELRDALR 360
DB 301 HLSTLPAARFNSGLQTLGTRNPLRSALPRGVQGLRELRLVGLHTNALAELRDALR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
DB 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389

RESULT 3
GPV_RAT STANDARD; PRT; 567 AA.
AC O08770;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Platelet glycoprotein V precursor (GPV) (CD42D).
GN Name:Gp5;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=97275136; PubMed=9129030;
RA Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S., Grunert P.,
RA Loew D., van Dorsselaer A., Cazenave J.-P., Lanza F.;
RT "Gene cloning of rat and mouse platelet glycoprotein V. Identification
RT of megakaryocyte-specific promoters and demonstration of functional
RT thrombin cleavage."
RL Blood 89:3253-3262 (1997).
CC -1- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand
CC factor receptor and mediates von Willebrand factor-dependent
CC platelet adhesion to blood vessels. The adhesion of platelets to
CC injured vascular surfaces in the arterial circulation is a

```

```

CC critical initiating event in hemostasis (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z69594; CAA93440.1; -.
CC HSSP; Q9BZR6; LOZN.
CC RGD; 2724; Gp5.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000483; LRR_Cterm.
CC InterPro: IPR000372; LRR_Nterm.
CC InterPro: IPR003591; LRR_Typ.
CC Pfam; PF00560; LRR; 14.
CC PRINTS; P00019; LEURICHRPT.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00362; LRR_Typ; 10.
CC Blood coagulation; Cell adhesion; Glycoprotein; Leucine-rich repeat;
CC Platelet; Repeat; Signal; Transmembrane.
CC SIGNAL 1 16 Potential.
CC CHAIN 17 567 Platelet glycoprotein V.
CC DOMAIN 17 522 Extracellular (Potential).
CC TRANSMEM 523 543 Potential.
CC DOMAIN 544 567 Cytoplasmic (Potential).
CC REPEAT 73 96 LRR 1.
CC REPEAT 97 120 LRR 2.
CC REPEAT 122 144 LRR 3.
CC REPEAT 145 168 LRR 4.
CC REPEAT 169 192 LRR 5.
CC REPEAT 194 216 LRR 6.
CC REPEAT 217 240 LRR 7.
CC REPEAT 241 264 LRR 8.
CC REPEAT 266 288 LRR 9.
CC REPEAT 289 312 LRR 10.
CC REPEAT 314 337 LRR 11.
CC REPEAT 338 361 LRR 12.
CC REPEAT 362 385 LRR 13.
CC REPEAT 387 409 LRR 14.
CC CARBOHYD 51 51 N-linked (GlcNAc. . .) (Potential).
CC CARBOHYD 181 181 N-linked (GlcNAc. . .) (Potential).
CC CARBOHYD 243 243 N-linked (GlcNAc. . .) (Potential).
CC CARBOHYD 298 298 N-linked (GlcNAc. . .) (Potential).
CC CARBOHYD 312 312 N-linked (GlcNAc. . .) (Potential).
CC CARBOHYD 385 385 N-linked (GlcNAc. . .) (Potential).
CC CARBOHYD 498 498 N-linked (GlcNAc. . .) (Potential).
SQ SEQUENCE 567 AA; 63344 MW; CA10708E0D03707F CRC64;

Query Match 90.2%; Score 1749; DB 1; Length 567;
Best Local Similarity 89.2%; Pred. No. 1.3e-109;
Matches 347; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPKCKCVRDAAQCSCGSAHAIAELGLPTNLTHILLFRM 60
DB 1 MLRSALLSAVLPLLRQAPFPCKCKCVRDAAQCSCGSAHAIAELGLPTNLTHILLFRM 60
QY 61 DOGILRNHSPSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTKNLSRLPRAILDK 120
DB 61 DOGILRNHSPSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTKNLSRLPRAILDK 120
QY 121 MVLEQFLDHNALRDLDQNLFOQLRNQLGELGNQNLQSLFPLNLFSSRELKLLDLSRN 180
DB 121 MVLEQFLDHNALRDLDQNLFOQLRNQLGELGNQNLQSLFPLNLFSSRELKLLDLSRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNIGALTEURLERNHLRSVAPGAFDR 240
DB 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNIGALTEURLERNHLRSVAPGAFDR 240

```

QY 241 LGNLSLTLGNLLESLPPALFLHVSVSRLTLFENPLEELPDVLFGEWAGRELWLNGT 300
 DB 241 LGNLSLTLGNLLESLPPALFLHVSWSRLTLFENPLEELPDVLFGEWAGRELWLNGT 300
 QY 301 HLSTLPAAPRNLSGLTGLTRNPLSALPRGVFGIENLRVLAHTNALAELEDDALR 360
 DB 301 HLSTLPAAPRNLSGLTGLTRNPLSALPRGVFGIENLRVLAHTNALAELEDDALR 360
 QY 361 GLGHLROVSURHNLRLALPRLFRNLSSL 389
 DB 361 GLGHLROVSURHNLRLALPRLFRNLSSL 389

RESULT 4

GPV_HUMAN
 ID - GPV_HUMAN STANDARD; PRT; 560 AA.
 AC P40197;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, last sequence update)
 DT 05-JUL-2004 (Rel. 44, last annotation update)
 DE Platelet glycoprotein V precursor (GPV) (CD42D).
 GN Name=GP5;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=93391348; PubMed=7690959;
 RA Hickey M.J., Hagen F.S., Yagi M., Roth G.J.;
 RT "Human platelet glycoprotein V: characterization of the polypeptide
 RT and the related Ib-V-IX receptor system of adhesive, leucine-rich
 RT glycoproteins."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8327-8331(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Platelet;
 RX MEDLINE=94012616; PubMed=8407908;
 RA Lanza F., Morales M., de la Salle C., Cazenave J.-P., Clemenston K.J.,
 RA Shimomura T., Phillips D.R.;
 RT "Cloning and characterization of the gene encoding the human platelet
 RT glycoprotein V. A member of the leucine-rich glycoprotein family
 RT cleaved during thrombin-induced platelet activation."
 RL J. Biol. Chem. 268:20801-20807(1993).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Platelet;
 RX MEDLINE=90275263; PubMed=2350580;
 RA Shimomura T., Fujimura K., Maehama S., Takemoto M., Oda K.,
 RA Fujimoto T., Oyama R., Suzuki M., Ichiara-Tanaka K., Titani K.,
 RA Kuramoto A.;
 RT "Rapid purification and characterization of human platelet
 RT glycoprotein V: the amino acid sequence contains leucine-rich
 RT repetitive modules as in glycoprotein Ib."
 RL Blood 75:2349-2356(1990).
 RN [4]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Platelet;
 RX MEDLINE=90321220; PubMed=2372284;
 RA Roth G.J., Church T.A., McMullen B.A., Williams S.A.;
 RT "Human platelet glycoprotein V: a surface leucine-rich glycoprotein
 RT related to adhesion."
 RL Biochem. Biophys. Res. Commun. 170:153-161(1990).
 CC -1- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand
 CC factor receptor and mediates von Willebrand factor-dependent
 CC platelet adhesion to blood vessels. The adhesion of platelets to
 CC injured vascular surfaces in the arterial circulation is a
 CC critical initiating event in hemostasis.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Platelets and megakaryocytes.
 CC -1- PTM: The N-terminus is blocked.

CC -1- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L11238; AAA03069.1; -;
 DR EMBL; Z33051; CA80637.1; -;
 DR FIR; A48030; A60164.
 DR HSSP; P07359; IM02.
 DR Genew; HGNC:4443; GP5.
 DR MIM; 173511; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR; 14.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR Blood coagulation; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Leucine-rich repeat; Platelet; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 16 Potential.
 FT CHAIN 17 560 Platelet glycoprotein V.
 FT DOMAIN 17 523 Extracellular (Potential).
 FT TRANSMEM 524 544 Potential.
 FT DOMAIN 545 560 Cytoplasmic (Potential).
 FT REPEAT 73 96 LRR 1.
 FT REPEAT 97 120 LRR 2.
 FT REPEAT 122 144 LRR 3.
 FT REPEAT 145 168 LRR 4.
 FT REPEAT 169 192 LRR 5.
 FT REPEAT 194 216 LRR 6.
 FT REPEAT 217 240 LRR 7.
 FT REPEAT 241 264 LRR 8.
 FT REPEAT 266 289 LRR 9.
 FT REPEAT 289 312 LRR 10.
 FT REPEAT 314 337 LRR 11.
 FT REPEAT 338 361 LRR 12.
 FT REPEAT 362 385 LRR 13.
 FT REPEAT 386 409 LRR 14.
 FT CARBOHYD 51 51 N-linked (GlcNAc...).
 FT CARBOHYD 181 181 N-linked (GlcNAc...).
 FT CARBOHYD 243 243 N-linked (GlcNAc...). (Potential).
 FT CARBOHYD 267 267 N-linked (GlcNAc...). (Potential).
 FT CARBOHYD 298 298 N-linked (GlcNAc...).
 FT CARBOHYD 312 312 N-linked (GlcNAc...).
 FT CARBOHYD 385 385 N-linked (GlcNAc...).
 FT CARBOHYD 499 499 N-linked (GlcNAc...). (Potential).
 FT CONFLICT 73 74 MT -> TK (in Ref. 2).
 FT CONFLICT 109 109 K -> T (in Ref. 2).
 FT CONFLICT 130 130 D -> W (in Ref. 3).
 FT CONFLICT 136 138 GID -> PGG (in Ref. 3).
 FT CONFLICT 209 209 L -> I (in Ref. 2).
 FT CONFLICT 267 267 N -> H (in Ref. 2).
 FT CONFLICT 327 327 L -> I (in Ref. 2).
 FT CONFLICT 478 478 P -> G (in Ref. 2).
 FT CONFLICT 509 509 P -> D (in Ref. 2).
 SQ SEQUENCE 560 AA; 60959 MW; BICB04AF8AF7115 CRC64;
 Query Match 76.2%; Score 1477; DB 1; Length 560;
 Best Local Similarity 76.1%; Pred. No. 2.7e-91;
 Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;
 QY 1 MURSALLSAVALLRAPFPCKTKCKVDRDAACGGGVAHAEGLPTNLTHILLFRM 60
 DB 1 MLRGTLCAVLGLLRAPFPCKTKCKVDRDAACGGGVAHAEGLPTNLTHILLFRM 60
 QY 61 DQGLRNHSFSGMTVLQRLMLSDSHSAIDPGTGNLVLKLTLRNKRISLRPAIDLK 120

Db 61 GGVVLOSFGSMVQLQMLSDSHISAVAGTSDLIKLTLSRNKTHLFGALLDX 120
QY 121 MYLLEQLFDHNAIRDLQDNFQQLRNQLQELGNQNSFLPANLFSRLKLLDLRN 180
Db 121 MYLLEQLFDHNAIRDLQDNFQQLRNQLQELGNQNSFLPANLFSRLKLLDLRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLSNLGNLALTELRLERHLSRVAPGAFDR 240
Db 191 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLSNLGNLALTELRLERHLSRVAPGAFDR 240
QY 241 LGNLSSTLSCNLLLESPPALFLHVSSVSRLTFLFENPLEELPDVLFGEAGLRELWNGT 300
Db 241 LGNLSSTLSCNLLLESPPALFLHVSSVSRLTFLFENPLEELPDVLFGEAGLRELWNGT 300
QY 301 HLSTLPAARFNLSGLQTLGTRPRLSALPRGVFGQLRRLVLAHTNALAEHRLDRLR 360
Db 301 QLRTPAAAFNRLGRRLVGLTSPRLSALPQGAQFGLGQLVLAHNSGLTALPDGLLR 360
QY 361 GLGHLRQVSLRHLRLRALPRTLFRNLSSL 389
Db 361 GLGHLRQVSLRHLRLRALPRTLFRNLSSL 389

RESULT 5

Q6DCV7 PRELIMINARY; PRT; 637 AA.
AC Q6DCV7
DT 01-OCT-2004 (Tremblrel. 28, Created)
DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=9355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Baha S.S., Loquellano N.A., Peters G.J., Carninci P., Frange C.,
RA Borak S., McEwan P.J., McKernan K.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;

RA Klein S., Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC077882; AAH77882.1; --
KW Hypothetical protein.
SQ SEQUENCE 637 AA; 72764 MW; 2D174B8895789F23 CRC64;
Query Match 40.4%; Score 782; DB 2; Length 637;
Best Local Similarity 45.3%; Pred. No. 2.3e-44;
Matches 167; Conservative 58; Mismatches 144; Indels 0; Gaps 0;
QY 21 CPKTCVKVVRDAACQSGGSAHIAELGLPTNLTHILLFRMDQGLRNHSGMTVLQRLM 80
Db 18 CPTLCTCRKDAVFCQSPKIDIGSLLPSNFTYIHINTLATETDKSFGNMPITLRLR 77
QY 81 LSDSHISALDPGTNDLVKLTRLTRNKIKSLPRAILDKMVLLEQLFDHNAIRDLQDN 140
Db 78 LEDSRITFTTRDAFKSLPQLKSLTNKLETPAGVFDLSFLYQLQFIGNVHLSLHPN 137
QY 141 LFQQLRNQLQELGNQNSFLPANLFSRLKLLDLRNLTLPKGLLGAQVKLEKLL 200
Db 138 LFCCQLHKLKELINRNQTLSPNELRLNLITLNLGRNKISHLPVIFSSLTCLKLH 197
QY 201 LYSNQLTSVDSGLSNLGNLALTELRLERHLSRVAPGAFDRNLSSLTSGNLLSPLPA 260
Db 198 LYENQLLITSSAFNNLGLLELALYSNLSIQSIADAFHPLPKLRLNLSXNKLHFLPYG 257
QY 261 LFLHVSSVSRLTFLFENPLEELPDVLFGEAGLRELWNGTHTLSTLPAARFNLSGLQTLG 320
Db 258 LFLHLPQLSVLTLYDNPILKELPDVIFGKVENLTSJLVYDTHLATIPNFVFCNLQLLV 317
QY 321 LTRNPRLSALPRGVFGQLRRLVLAHTNALAEHRLDRLRGLHRLQVSLRHLRLRALPR 380
Db 318 LTRNPRLSALPRGVFGQLRRLVLAHTNALAEHRLDRLRGLHRLQVSLRHLRLRALPR 380
QY 381 LTRNPRLSALPRGVFGQLRRLVLAHTNALAEHRLDRLRGLHRLQVSLRHLRLRALPR 377
Db 378 NMVFNLSNL 386

RESULT 6

LR15_HUMAN
ID LR15_HUMAN STANDARD; PRT; 581 AA.
AC Q8TF66;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Leucine-rich repeat-containing protein 15 precursor (hLib).
GN Name=LRC15; Synonyms=LiB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21645900; PubMed=11795964; DOI=10.1006/bbrc.2001.6272;
RA Satoh K., Hara M., Yokota H.;
RT "A novel member of the leucine-rich repeat superfamily induced in rat
RT astrocytes by beta-amyloid";
RL Biochem. Biophys. Res. Commun. 290:756-762(2002).
CC -! SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -! TISSUE SPECIFICITY: Brain and placenta.
CC -! SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB071037; BAB84587.1; --

DR HSSP; Q9BZR6; 10ZN.
 DR Genew; HGNC:20818; LRRCL5.
 DR InterPro; IPR011061; Antihaemostatic.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR; 14.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_Typ; 10.
 KW Leucine-rich repeat; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 21
 FT CHAIN 22 581
 FT
 FT DOMAIN 22 538
 FT TRANSMEM 539 559
 FT DOMAIN 560 581
 FT REPEAT 51 75
 FT REPEAT 76 99
 FT REPEAT 100 123
 FT REPEAT 125 147
 FT REPEAT 148 171
 FT REPEAT 173 195
 FT REPEAT 196 219
 FT REPEAT 221 243
 FT REPEAT 244 267
 FT REPEAT 269 291
 FT REPEAT 292 315
 FT REPEAT 317 339
 FT REPEAT 340 363
 FT REPEAT 364 387
 FT REPEAT 389 411
 FT CARBOHYD 75 75
 FT CARBOHYD 369 369
 FT N-linked (GlcNAc...) (Potential).
 FT N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 581 AA; 64396 MW; 1F381485BD2CCB54 CRC64;

Query Match 31.3%; Score 607; DB 1; Length 581;
 Best Local Similarity 37.7%; Pred. No. 1.3e-32;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVVRDAAQCQSGSVAHIAELGLPTNLTTHILLFRMDQGLRNHSGFMTVLQRLM 80
 Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAWSLQILNTHITELNESPFNLISALIALR 83
 QY 81 LSDSHISAIDPGTFENDLVKLTIRLNKISRLPRAILDKMWLLQLFQGLDLSLESLSSNQLLIQPA 140
 Db 84 IEKNELSRITPGAFRNLSLRYSLANNKLVLPGLFQGLDLSLESLSSNQLLIQPA 143
 QY 141 LFOQLRNQLGELGNQNLQSLFPANLPSLSRELKLLDLSRNNTLHPKGLLGAQVKLEKL 200
 Db 144 HFSQCSNLKELQHLGNHLEYIPDGFHLVGLTKLNIGKNSLTHISPRVQHLGNLQVLR 203
 QY 201 LYSNQLTSVDSGLLSNLGALTLELRNHLRSVAPGAFDRLGNLSLITLSGNLESPPA 260
 Db 204 LYENRLTIPMGTFDGLVNLQELALQOQIGLSLPGLFHNNHNLQRLYLSNNHISQLPPS 263
 QY 261 LFLHVSSVRLTLPENPLEELPDVLFGEAGLRELWNGTHLSTLPAAAFRNLSGLQTLG 320
 Db 264 IFMQLPQNLRLTLFGNSLKLSLGI FGPMFNLRELWLYDNLHISLSPDNVFSNLRQLQVLI 323
 QY 321 LTRNPRLSALPRGVFGQLRELRLVLAHTNALAEIRDALRGHLRGVSLRHNRRLALPR 380
 Db 324 LSRN-QISFTSPGAFNGLTRELRLSHLTHNALQDLQDGNVFRMLANLQNLISLQNNRLQLPG 382
 QY 381 TLFRNLSSL 389
 Db 383 NIFANVNGL 391

RESULT 7
 QYTRIN7

ID QYTRIN7 PRELIMINARY; PRT; 581 AA.
 AC QYTRIN7;
 DT 01-MAR-2004 (TRENBLrel. 26, Created)
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE LRRCL5.
 GN Name=LRRCL5;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833479; PubMed=12923058;
 RA Reynolds P.A., Smolen G.A., Palmer R.E., Sgroi D., Yajnik V.,
 RA Gerald W.J., Haber D.A.;
 RT "Identification of a DNA-binding site and transcriptional target for
 RT the EWS-WT1(+KTS) oncoprotein.";
 RL Genes Dev. 17:2094-2107(2003).
 CC -1- MISCELLANEOUS: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
 DR EMBL; BK001325; DAA01740.1; ..
 DR InterPro; IPR001611; LRR_Typ.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR; 14.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SEQUENCE 581 AA; 64396 MW; 1F381485BD2CCB54 CRC64;

Query Match 31.3%; Score 607; DB 2; Length 581;
 Best Local Similarity 37.7%; Pred. No. 1.3e-32;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKCKCVVRDAAQCQSGSVAHIAELGLPTNLTTHILLFRMDQGLRNHSGFMTVLQRLM 80
 Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAWSLQILNTHITELNESPFNLISALIALR 83
 QY 81 LSDSHISAIDPGTFENDLVKLTIRLNKISRLPRAILDKMWLLQLFQGLDLSLESLSSNQLLIQPA 140
 Db 84 IEKNELSRITPGAFRNLSLRYSLANNKLVLPGLFQGLDLSLESLSSNQLLIQPA 143
 QY 141 LFOQLRNQLGELGNQNLQSLFPANLPSLSRELKLLDLSRNNTLHPKGLLGAQVKLEKL 200
 Db 144 HFSQCSNLKELQHLGNHLEYIPDGFHLVGLTKLNIGKNSLTHISPRVQHLGNLQVLR 203
 QY 201 LYSNQLTSVDSGLLSNLGALTLELRNHLRSVAPGAFDRLGNLSLITLSGNLESPPA 260
 Db 204 LYENRLTIPMGTFDGLVNLQELALQOQIGLSLPGLFHNNHNLQRLYLSNNHISQLPPS 263
 QY 261 LFLHVSSVRLTLPENPLEELPDVLFGEAGLRELWNGTHLSTLPAAAFRNLSGLQTLG 320
 Db 264 IFMQLPQNLRLTLFGNSLKLSLGI FGPMFNLRELWLYDNLHISLSPDNVFSNLRQLQVLI 323
 QY 321 LTRNPRLSALPRGVFGQLRELRLVLAHTNALAEIRDALRGHLRGVSLRHNRRLALPR 380
 Db 324 LSRN-QISFTSPGAFNGLTRELRLSHLTHNALQDLQDGNVFRMLANLQNLISLQNNRLQLPG 382
 QY 381 TLFRNLSSL 389
 Db 383 NIFANVNGL 391

RESULT 8
 QYTRIN7 PRELIMINARY; PRT; 579 AA.
 AC QYTRIN7;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Lrrc15 protein.
 GN Name=Lrrc15;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RA Strausberg R.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC050345; AAH50245.1;
DR HSSP; Q9BZR6; 102N.
DR MGD; MGI:1921738; Lrrc15.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 14.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 11.
SQ SEQUENCE 579 AA; 64176 MW; 7841957F6FB643D3 CRC64;

Query Match 31.08; Score 600; DB 2; Length 579;
Best Local Similarity 36.94; Pred. No. 3.8e-32;
Matches 137; Conservative 62; Mismatches 170; Indels 2; Gaps 2;

Qy 19 FPCPKTCVVRDAAQCSCGSGVAHIAELGLPTNLTHILLFRMDQGLRNHFSGMTVLQR 78
Db 23 YGCPSECTCSRAQVECTGAQIVAMPS-PLPWNAMSLQILNTHITELPEDKFNISALIA 81
Qy 79 LMLSDSHISAIDPGTFNDLVKLTLRITRNKISRLPRAILDKNVLEQLFDHNAIRDLD 138
Db 82 LKMEKELANIMFCAFNLSGRLSLANNKLNFLVFDVNNLETLLSNNQLVQIQ 141
Qy 139 QNLFOQLRNQELGLNQNLSPANLFFSLRELKLLDLRNLTLPKGLLCAQVKLEK 198
Db 142 PAQFSQESNKLKQLVGNLLEYIPGVDFHLVGLTKLNGNGFTHLSFVQHLGNLQV 201
Qy 199 LLLYSNQLTSVDSGLSNIGALTELRLRNHLSVAPGAFDRIGNSSITLSCNLESIP 258
Db 202 LRLYENRLSDIPMGTTFDALGNLQELALQENQIGTSLSPGLFHNRLNRLYLSNNHISLP 261
Qy 259 PALFLHVSYSRLTLFENPLEEYPDVLFCEMAGRLWLNGLTHLSLTPAAAFNLSGLQT 318
Db 262 PGIFMQPHLNKLTLPFNSLKEUSPGVFGMPMLRELWLNHITSLPDAFSLNQLQV 321
Qy 319 LGLTRNFRISALPRGVFQGLRELRLVLAHTNALAEIRDDALRGLHRLQVSLRHRRLAL 378
Db 322 LILSHN-QLSYISPGAFNGTLNRELSLHTNALQDLQDGNVFRSLANRLNVSQNRLRLQ 380
Qy 379 PRTLFRNLSSL 389
```

```

Db 381 FGSIFANVGL 391
ID LR15 RAT STANDARD; PRT; 578 AA.
AC Q8R5M3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Leucine-rich repeat-containing protein 15 precursor (rLib).
GN Name=Lrrc15; Synonyms=Lib;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Fetal brain;
RX MEDLINE=21645900; PubMed=11785964; DOI=10.1006/bbrc.2001.6272;
RA Satoh K., Hata M., Yokota H.;
RT "A novel member of the leucine-rich repeat superfamily induced in rat
RT astrocytes by beta-amyloid.";
RL Biochem. Biophys. Res. Commun. 290:756-762(2002).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- INDUCTION: By beta-amyloid.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB071036; BAB84586.1;
CC HSSP; Q9BZR6; IP8T.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000483; LRR_Cterm.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003591; LRR_typ.
CC Pfam; PF00560; LRR; 14.
CC PRINTS; PR00019; LEURICHRPT.
CC SMART; SM00082; LRRCT; 1.
CC SMART; SM00013; LRRNT; 1.
CC SMART; SM00369; LRR_TYP; 11.
KW Leucine-rich repeat; Repeat; Signal; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 578
FT
FT DOMAIN 22 535
FT Extracellular (Potential).
FT TRANSMEM 536 556
FT DOMAIN 557 578
FT Cytoplasmic (Potential).
FT REPEAT 51 75
FT LRR 1.
FT REPEAT 76 99
FT LRR 2.
FT REPEAT 100 123
FT LRR 3.
FT REPEAT 125 147
FT LRR 4.
FT REPEAT 148 171
FT LRR 5.
FT REPEAT 173 195
FT LRR 6.
FT REPEAT 196 219
FT LRR 7.
FT REPEAT 221 243
FT LRR 8.
FT REPEAT 244 267
FT LRR 9.
FT REPEAT 269 291
FT LRR 10.
FT REPEAT 292 315
FT LRR 11.
FT REPEAT 316 339
FT LRR 12.
FT REPEAT 340 363
FT LRR 13.
FT REPEAT 364 387
FT LRR 14.
FT REPEAT 389 411
FT LRR 15.
FT CARBOHYD 75 75
FT N-linked (GlcNAc...) (Potential).
FT CARBOHYD 369 369
FT N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 578 AA; 64127 MW; 991BD057F5912591 CRC64;
```



```
QY 319 LCLTNPRLSALPRGVFOGLRELRLVIALHTNALAEALRDLALRGHLRQVSLRHLNRRLAL 378
Db 322 LILSHN-QLSVISPGAFNGLNLRLSLHTNALQDLQDNVPSLANL---GTSHSRITAF 377
QY 379 PRTL 382
Db 378 DSSL 381

RESULT 11
Q8R113 PRELIMINARY; PRT; 562 AA.
AC Q8R113;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE 1300018K11Rik protein (Fragment).
GN Name=1300018K11Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025836; AAH25836.1; -.
DR HSPF; P07359; IWOZ.
DR MGD; MGI:1919006; 1300018K11Rik.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Cyp.
DR Pfam; PF00560; LRR_12.
DR Pfam; PF01462; LRRN1; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 8.
FT NON TER
FT 1
SQ SEQUENCE 562 AA; 61993 MW; C086F0BCFAEF09AE CRC64;
Query Match 26.0%; Score 503.5; DB 2; Length 562;
Best Local Similarity 34.6%; Pred. No. 1.2e-25;
Matches 129; Conservative 69; Mismatches 172; Indels 3; Gaps 3;
QY 11 LALLRAQFFPCPXTCKVCVWDAAQCSCGSAHIAELGLPLNTHILLFRMDQGLRNHSF 70
```

```
Db 27 LULLARLTQPCPGVGCFCGRVFP-CSDEQLADIPP-DIPPHITDIVFVETATTVTRAF 84
QY 71 SGMTVLQRLMDSHSISAIIDPOTFNDLVKLTIRTRNKISRLPRAILDKNVLLQLFLD 130
Db 85 SGSPNLTQVFLNTQVRHLEPDAFGGLPRLOLEITGSPVSNLSAHIFSLSLEKLTLD 144
QY 131 HNALRDLQNLQFQRLNQLQELGNQNLQFLPANLFFSLRELKLLDLGRNNLTHLPKGLL 190
Db 145 FDELACLPEDLPCHMDILESLOLQNLQRTLPGRLFQSLRDLRTLNLAQNLTLQPKGAF 204
QY 191 GAQVKLEKLLYSNQLTSVDSGLLSNGLALTELRNLRLSRVAPGADRGLNLSLTL 250
Db 205 QSLTGLQMLKLSNNMLARPELAGLSLQSLQELFDGNAITELSPHLSQFLSLEMLWQ 264
QY 251 GNLLSPPALFLHVSSVRLTLFENPLEELPDVLFGEWAGRELWANGTHLSTLPAAAF 310
Db 265 HNAICHLPVSLFSSLHNLFTLSKONALRTLDEGLFAHQGLLHLSLSYNQLETIPEGAF 324
QY 311 RNLSGLQTLGLTRNPRLSALPRGVFGGLRELRLVLAHTNALAEALRDLALRGHLRQVSL 370
Db 325 TNLRLVSLTLGSHN-AITDLPHEVFRNLQVLKLSLDSNNLTALHPALFHLNLSRLQLNL 383
QY 371 RHNRLALPRTL 383
Db 384 SRNQLTTPGGIF 396

RESULT 12
Q9DBB9 PRELIMINARY; PRT; 570 AA.
AC Q9DBB9;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Mus musculus adult male liver cDNA, RIKEN full-length enriched
DE library, clone:1300018K11 product weakly similar to CARBOXYPEPTIDASE N
DE 93 kDa CHAIN (CARBOXYPEPTIDASE N REGULATORY SUBUNIT).
GN Name=1300018K11Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RN
```

Genome Res. 10:1617-1630(2000).

[5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Kitsuai T., Tashiro H., Itoh M.,
 RA Konno H., Akiyama J., Nishi K., Hazama M., Nishine T., Harada A.,
 RA Sumi N., Ishii Y., Nakamura S., Ikegami T., Kashiwagi K.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirakawa T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sorabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Ito T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.,
 RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AK005049; BAB23775.1; -
 DR HSSP; P07359; 1M0Z.
 DR MGD; MGI:1913006; 1300018K11Rik.
 DR GO; GO:0004180; F-carboxypeptidase activity; IEA.
 DR InterPro; IPR001611; LRR_Cterm.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR; 12.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_Typ; 8.
 KW Carboxypeptidase.
 SQ SEQUENCE 570 AA; 62889 MW; 634CEE2E5F467C7F CRC64;

Query Match 26.0%; Score 503.5; DB 2; Length 570;
 Best Local Similarity 34.6%; Pred. No. 1.2e-25;
 Matches 129; Conservative 69; Mismatches 172; Indels 3; Gaps 3;

QY 11 LALLRAQPPCPKTKCVVRDAACQSGGVVAHIAELGLPTNTHILLFRMDQGLRNHSHF 70
 DB 35 LLLLRATQCPVGCDCFCGEVFP-CSDQELADIPP-DIPPHITDVFVETAFVTRAF 92
 QY 71 SGMTVQLRLMDSHSAIDPPGFNDLVKLTIRLRNKRISRLPRAILDQKVLLEOLF 130
 DB 93 SGSPNLKVVFLNTQVRHLEPDAGGLPRLODLEITGSPVNSLHSAHISLSSLEKLTLD 152
 QY 131 HNALRDQNLQQLRNQLGLNQLSFLPANLPSSRLRELKLDLSRNNLTHLPKGLL 190
 DB 153 FDLRAGLPEDFCHMDILESLOQGNQLRTLPGRFQSLRDLRTLNLAQNLTLPKGA 212
 QY 191 GAQVKEKLLIYNQLTSDVDSGLLSNLGALTLELRNHLRSVAPGADRLGNLSLTLIS 250
 DB 213 QSLTGLQMLKNNMLARPEGALSGLSLQELFLDGNATIELSPHLSQLFSLEMLW 272
 QY 251 GNLESIPPALFLHVSSELTLTFENPLELDDVLPFGVAGRELWNGTHLSTLPAAAF 310
 DB 273 HNAICHLPVSLFSLSHNLTLFLSKDNALRTLPEGUFAHQGLHLSLSTNQLEIPEGAF 332
 QY 311 RNLSGLQTLGLTRNPRISALPRGVFGQLRELRLVLAHTNALAEALRDLRGLHRLQVSL 370

333 TNLSRLVSLTSLSEN-AITDLPEHVRNLEQVLKSLSDSNLTLHPALPHNLRLQLLNL 391

QY 371 RHNEPALPRTLF 383
 DB 392 SRNQLTTLPGGIF 404

RESULT 13
 Q8NSV4 PRELIMINARY; PRT; 557 AA.
 AC Q8NSV4;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
 DE Similar to RIKEN CDNA 1300018K11 gene (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR BCO31569; RAH31569.1; -
 DR HSSP; Q8ZB36; 1P9T.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR Pfam; PF00560; LRR; 13.
 DR Pfam; PF01462; LRRCT; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_Typ; 10.
 FT NON TER 1
 SQ SEQUENCE 557 AA; 61782 MW; A65BFC340CC76563 CRC64;

Query Match 23.0%; Score 445; DB 2; Length 557;
 Best Local Similarity 33.1%; Pred. No. 1e-21;
 Matches 125; Conservative 54; Mismatches 173; Indels 26; Gaps 4;

QY 1 MLRSALLSAVALLRAQPP-PCPKTKCVVRDAACQSGGVVAHI----- 43
 DB 13 MLPGAWLLWTSLLILARPAQPCMGCDQVQEVF-CSDEELATVPLDIPPYTKNIIFVET 71
 QY 44 -----AELGLPTNTHILLFRMDQGLRNHSHFSGMTVQLRLMDSHSAIDPPGFND 96
 DB 72 STTTLETRAFSGSNPLTKVFLNTQLCOFRPDAGGLPRLEDEVTGSSFLNISTNFSN 131
 QY 97 LVKLKTLRLTRNKRISRLPRAILDQKVLLEOLFDPHNLALRDQNLQQLRNQLGLNQN 156
 DB 132 LTLGKLTINFNMLPEGLFOHLALESILHQLGQLOALPRFLQPLTHLTLNLAQN 191
 QY 157 QLSFELPANLPSSRLRELKLDLSRNNLTHLPKGLLAGQVKEKLLIYNQLTSDVDSGLLS 216
 DB 192 LLAQVPEELFPHLTSLQTLKLSNNALSGLPQGVFGKLSQELFLDSNNISELPQVFSQ 251
 QY 217 LGALTLELRNHLRSVAPGADRLGNLSLTLSGNLSLPPALFLHVSSELTLTFEN 276
 DB 252 LFCLELRLQRNAITHLPLSLFASLGNLTFLSLQNNMLRVLPAGLFAHTPCVLGLSLT 311
 QY 277 PLEELPDVLPFGSMAGRELWNGTHLSTLPAAAFNLISGLQTLGLTRNPRISALPRGVFG 336
 DB 312 QLETVAEGTFAHLSNLSLMSYNATHLPAGIFRDLLEELVKLYLGSN-NLTALHPALFQ 370
 QY 337 GRELRLVLAHTNALAE 354
 DB 371 NL-SKLELLSLSKNQLTTL 388


```

RESULT 14
ID Q86SU4 PRELIMINARY; PRT; 560 AA.
AC Q86SU4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to RIKEN cDNA 1300018K11 gene (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC042334; AA042334.1; -.
DR HSSP; Q9BZR6; 10ZN.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR; 13.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHPT.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 10.
FT NON TER
FT SEQUENCE 560 AA; 62196 MW; CABFEBCE1E69B0934 CRC64;
Query Match 23.08; Score 445; DB 2; Length 560;
Best Local Similarity 33.18; Pred. No. 1e-21;
Matches 128; Conservative 58; Mismatches 173; Indels 26; Gaps 4;
QY 1 MLRSALLSAVLALLRAQPF-PCPKTKCVVRDAAQCQSGGSAHH----- 43
Db 16 MLPQAWLLWTSLLLRPAQPCPMGDCDFVQEVF-CSDEBLATVPLDIPYTKNIIFVET 74
QY 44 -----AELGLPTNLTHILLFMDQGLRNHFSGMVLRQLMLSDSHISAIDPTFND 96
Db 75 SFTLETRAGSNPNKTKVFLNQLCQFDPDAGFLPRLEDLEVGGSLNLTNIFGN 134
QY 97 LVKLKTLRLTRNKISRLPRAILDKVLLLEQLFDHNAIRDLDQNLFOQLRNQLQELGNQ 156
Db 135 LTSIGKLTNFMNLEALPEGLFOHLLAALSLHLCQNLQALPRLLQPLTHLKTNLQ 194
QY 157 QLSFLPANLSSIREKLKLDLSRNLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLN 216
Db 195 LLAQLPEELPHLTSLSQTLKSNALSGLPQGVFGKLSLQELFLDSNNISELPPQVFSQ 254
QY 217 LGALTELRLEARNLRSVAPGAPDRGLNLSLTLGSLNLESLEPALFLHVSYSRLTLFEN 276
Db 255 LFCLERLWLRNATHLPLSIFASLGNLFLSLQWNLRVLPAGLFAHTPCVLGSLTN 314
QY 277 PLEELPVLFGEMAGRELWNLGTHSLTPAAAFRLNLSGLQTLGLTRNPRLSALPGEVFQ 336
Db 315 QLETVAEGTFAHLSNLSRLMLSNAYTHLPAGIFRDLLELVKLYLGSN-NLTALHPALFQ 373
QY 337 GLRELRLVLAHTWALAEI 354
Db 374 NLSKLELLSLSKNQLTTL 391
RESULT 15
ALS HUMAN
ID ALS HUMAN STANDARD; PRT; 605 AA.
AC P35858;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

```

```

DE Insulin-like growth factor binding protein complex acid labile chain
DE Precursor (ALS).
DE Name=IGFALS; Synonyms=ALS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92357025; PubMed=1379671;
RA Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RT "Structure and functional expression of the acid-labile subunit of the
RL insulin-like growth factor-binding protein complex.";
RL Mol. Endocrinol. 6:870-876(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20114388; PubMed=10650966;
RA Suwanichkul A., Boisclair Y.R., Olney R.C., Durham S.K., Powell D.R.;
RT "Conservation of a growth hormone-responsive promoter element in the
RL human and mouse acid-labile subunit genes.";
RL Endocrinology 141:833-838(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 28-35.
RX MEDLINE=89308584; PubMed=2473065;
RA Baxter R.C., Martin J.L., Beniac V.A.;
RT "High molecular weight insulin-like growth factor binding protein
RL complex. Purification and properties of the acid-labile subunit from
RT human serum.";
RL J. Biol. Chem. 264:11843-11848(1989).
CC -!- FUNCTION: Involved in protein-protein interactions that result in
CC protein complexes, receptor-ligand binding or cell adhesion.
CC -!- SUBUNIT: Forms a ternary complex of about 140 to 150 kDa with IGF-
CC I or IGF-II and IGBP-3.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Contains 20 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M86826; AAA36047.1; -.
CC EMBL; AF192554; AAF06774.1; -.
CC EMBL; AL031724; CAC36078.1; -.
CC PIR; A41915; A41915.
CC HSP; P07359; IMOZ.
CC Genew; HGNC:5468; IGFALS.
CC MIM; 601489; -.
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0005625; C:soluble fraction; TAS.
CC GO; GO:0005520; F:insulin-like growth factor binding; TAS.
CC GO; GO:0007185; P:signal transduction; TAS.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR000372; LRR_Nterm.
CC InterPro; IPR003591; LRR_Typ.
CC Pfam; PF00560; LRR; 19.
CC Pfam; PF01462; LRRNT; 1.
CC PRINTS; PR00019; LEURICHPT.
CC Cell adhesion; Direct protein sequencing; Glycoprotein;
KW Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605 Insulin-like growth factor binding
FT protein complex acid labile chain.
FT REPEAT 53 73 LRR 1.

```

```
FT REPEAT 74 96 LRR 2.
FT REPEAT 96 120 LRR 3.
FT REPEAT 121 144 LRR 4.
FT REPEAT 145 168 LRR 5.
FT REPEAT 170 192 LRR 6.
FT REPEAT 193 216 LRR 7.
FT REPEAT 217 240 LRR 8.
FT REPEAT 242 264 LRR 9.
FT REPEAT 265 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 337 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 480 LRR 18.
FT REPEAT 482 504 LRR 19.
FT REPEAT 505 530 LRR 20.
FT CARBOHYD 64 64 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 85 85 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 96 96 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 368 368 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 515 515 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 580 580 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 605 AA; 66034 MW; F6562A23CB918F6 CRC64;

Query Match 21.6%; Score 419.5; DB 1; Length 605;
Best Local Similarity 33.0%; Pred. No. 5.9e-20;
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

QY 38 GSAH---IAELGLPNTLHLLFMDDQILRNHSGMTVLQRLMLSDSHISAIDPCTF 94
Db 140 GTFAPTPALASGLSN---RSLRDEL-----FGLGSLWDLNCGWNSLAVLPDAF 190

QY 95 NDIVKLTLELRNKISRLPRAILDKRWLLQLFLEHNAIRDLDQNLFOQLRNLOELGLN 154
Db 191 RGLGSLRELVLACNRLAYLPALFSGLAELRELDLSRNALRAIKANVFVQLPRQKLYLD 250

QY 155 ONQLSFLPANLSSRELKLLDLSRNLTLPK---GLLGAQVKLEKLLYSNQLTSVD 210
Db 251 RNIIAIVAPGAFGLKALRWLDLSHNRVAGLLEDTFPGLLGURV---LRSHNAIASLR 306

QY 211 SGLLSNIGALTELELRNHLRSVAPGAFDRGLNLSLTLSGNLLSPPALFLHVSSYSR 270
Db 307 PRTFKDLHFEELQIGHNRIQLAERSFEGIQLEVLTLIDHNQLQEVKAGAFGLTNVAV 366

QY 271 LTLFENPLELPDLFEGMAGLEELWNGTHLSTLPAAFRNLGLQTLGTRN----- 324
Db 367 MNLSGNCLRLPEQVFRGLGKLSHLBSGSLGRIRPHFTTGLSGRLRELFKDNGLVGIE 426

QY 325 -----PRLSALPRGVFQGLRELVLAHTNNAELRDDLALRGHLRQ 367
Db 427 EQSLWGLAELELDLTSNQLTHPHLELFQGLKLEYLLLSRNRLAELPADALGPLQAFW 486

QY 368 VSLHNRRLRALPRLFNLSL 389
Db 487 LDVSHNRLEALPNSLAPLGR 508
```

Search completed: November 29, 2004, 13:10:59
Job time : 113 secs

Db 1 MLRGTLLCAVLGLLRAQPPCPCKCVFRDAAQCSGGDVARIASALGLPTNLTHTLLFGM 60
QY 61 DQILRNHFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK 120
Db 61 GRGVLSQSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK 120
QY 121 MVLLQQLFDHNAERLDONLFOQLRNQELGNQOLSPFANLSSRLKLLDLSRN 180
Db 121 MVLLQQLFDHNAERLDONLFOQLRNQELGNQOLSPFANLSSRLKLLDLSRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLRNHLSRVAPGAFDR 240
Db 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLRNHLSRVAPGAFDR 240
QY 241 LGNLSLTLSCNLLLESPPALFLHVSVSRITLTFENPLEELPDVLFGEWAGRLWNGT 300
Db 241 LGNLSLTLSCNLLLESPPALFLHVSVSRITLTFENPLEELPDVLFGEWAGRLWNGT 300
QY 301 HLSTLPAAAFRNLSGLQTLGTENPRLSALPRGVFGQLRELRLVLAHTNALAELRDDAIR 360
Db 301 QLRTPAAAFRNLSRLRYLGVTLSPRLSALPQAFQGLGELQVLAHNSGLTALPDGLLR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
Db 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389

RESULT 2

US-08-195-006-2
; Sequence 2, Application US/08195006
; Patent No. 6083688
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/195,006
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12418-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-195-006-2

Query Match 76.2%; Score 1477; DB 3; Length 560;
Best Local Similarity 76.1%; Pred. No. 1.9e-135;

Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;
QY 1 MLRSALLSAVLALLRAQPPCPCKCKVDRDAAQCSGGSVAHIAELGLPTNLTHTLLFGM 60
Db 1 MLRGTLLCAVLGLLRAQPPCPCKCKVDRDAAQCSGGDVARIASALGLPTNLTHTLLFGM 60
QY 61 DQILRNHFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK 120
Db 61 GRGVLSQSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK 120
QY 121 MVLLQQLFDHNAERLDONLFOQLRNQELGNQOLSPFANLSSRLKLLDLSRN 180
Db 121 MVLLQQLFDHNAERLDONLFOQLRNQELGNQOLSPFANLSSRLKLLDLSRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLRNHLSRVAPGAFDR 240
Db 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLRNHLSRVAPGAFDR 240
QY 241 LGNLSLTLSCNLLLESPPALFLHVSVSRITLTFENPLEELPDVLFGEWAGRLWNGT 300
Db 241 LGNLSLTLSCNLLLESPPALFLHVSVSRITLTFENPLEELPDVLFGEWAGRLWNGT 300
QY 301 HLSTLPAAAFRNLSGLQTLGTENPRLSALPRGVFGQLRELRLVLAHTNALAELRDDAIR 360
Db 301 QLRTPAAAFRNLSRLRYLGVTLSPRLSALPQAFQGLGELQVLAHNSGLTALPDGLLR 360
QY 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389
Db 361 GLGHLRQVSLRNHLRALPRTLFRNLSSL 389

RESULT 3

US-09-063-950-4
; Sequence 4, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-063-950-4

Query Match 76.2%; Score 1477; DB 3; Length 560;
Best Local Similarity 76.1%; Pred. No. 1.9e-135;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLALLRAQPPCPCKCKVDRDAAQCSGGSVAHIAELGLPTNLTHTLLFGM 60
Db 1 MLRGTLLCAVLGLLRAQPPCPCKCKVDRDAAQCSGGDVARIASALGLPTNLTHTLLFGM 60
QY 61 DQILRNHFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK 120
Db 61 GRGVLSQSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLTRLTRNKISRLPRAILDK 120
QY 121 MVLLQQLFDHNAERLDONLFOQLRNQELGNQOLSPFANLSSRLKLLDLSRN 180
Db 121 MVLLQQLFDHNAERLDONLFOQLRNQELGNQOLSPFANLSSRLKLLDLSRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLRNHLSRVAPGAFDR 240
Db 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTTELRLRNHLSRVAPGAFDR 240
QY 241 LGNLSLTLSCNLLLESPPALFLHVSVSRITLTFENPLEELPDVLFGEWAGRLWNGT 300
Db 241 LGNLSLTLSCNLLLESPPALFLHVSVSRITLTFENPLEELPDVLFGEWAGRLWNGT 300

QY 301 HLTSLPAAAFNRLSGLTGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360
DB 301 QLRTPAAAFNRLSRRLVGLTSLPRLSALPQGAFOGLGQVLAHSNGLTALPDGLLR 360
QY 361 GLGHLRQVSLRHRNRLRALPRTLFRNLSSL 389
DB 361 GLGKLRQVSLRHRNRLRALPRLFRNLSSL 389

RESULT 4

PCT-US94-07644A-2
; Sequence 2, Application PC/TUS9407644A
; GENERAL INFORMATION:
; APPLICANT: COR Therapeutics, Inc.
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07644A
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 012418-003000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-07644A-2

Query Match 76.2%; Score 1477; DB 5; Length 560;
Best Local Similarity 76.1%; Pred. No. 1.9e-135;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSALLSAVLAILRAQFPCKCKVDRDAQCSCGSAHIAELGLPTNLTHILLFRM 60
DB 1 MLRRTLCAVLGLLRQAFPCKCKVDRDAQCSCGSDVARISALGLPTNLTHILLFGM 60
QY 61 DQGLRNHSFSGMTVLQRLMSDSHISAIDPGTFNDLVKLTURLTRNKISRLPRAILDK 120
DB 61 GRGVLSQSFSGMTVLQRLMSDSHISAIVAGTFSDIILKLTLSNKITHILPGALLDK 120
QY 121 MVLEQLFLDHNAERDLDONLFOQLRNQLGELGNQOLSFPLNPFSSRELKLLDSRN 180
DB 121 MVLEQLFLDHNAERDLDONLFOQLRNQLGELGNQOLSFPLNPFSSRELKLLDSRN 180
QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNGLATLRLERNLHRSVAPGAFR 240
DB 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNGLATLRLERNLHRSVAPGAFR 240
QY 241 LGNISLTLNLSLESPPALFLHVSSVSKLTTFENPLEEPDVLFGEMAGRELWINGT 300
DB 241 LPNLSLTLNLSLESPPALFLHVSSVSKLTTFENPLEEPDVLFGEMAGRELWINGT 300
QY 301 HLTSLPAAAFNRLSGLTGLTRNPRLSALPRGVFQGLRELRLVLAHTNALAEIRDALR 360

DB 301 QLRTPAAAFNRLSRRLVGLTSLPRLSALPQGAFOGLGQVLAHSNGLTALPDGLLR 360
QY 361 GLGHLRQVSLRHRNRLRALPRTLFRNLSSL 389
DB 361 GLGKLRQVSLRHRNRLRALPRLFRNLSSL 389

RESULT 5

US-08-190-802A-49
; Sequence 49, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Insulin-like growth factor binding
; INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-190-802A-49

Query Match 21.6%; Score 419.5; DB 1; Length 605;
Best Local Similarity 33.0%; Pred. No. 3.7e-32;
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

QY 38 GSAH---IAGLGPNTLTHILLFRMDQGLRNHSFSGMTVLQRLMSDSHISAIDPGTF 94
DB 140 GFAHTPALASGLSNN---RLSRLEDGL-----FEGLSGLWDLNGLWNSLAVLPDAAF 190
QY 95 NDLVKLTURLTRNKISRLPRAILDKMLLEQLFLDHNAERDLDONLFOQLRNQLGELGN 154
DB 191 RGLGSLRELVLNGLAVLQPALFSGLAELRELDLSRNALRAIKANVFPVQLPRIOKLYLD 250
QY 155 QNQLSFLPANLPSSURELKLLDLSNNLTHLPK---GLLGAQVKLEKLLYSNQLTSVD 210
DB 251 RNLIAAVAPGAFGLKALRWLDLSHNRVAGLLEDTFGLGLRV---LRLSHNAIASLR 306
QY 211 SGLLSNGLATLRLERNLHRSVAPGAFDRGNLSLTLNLSLESPPALFLHVSSVSR 270
DB 307 PRTFKDLHLEQLGHNRIRQLAERSFEGLGQLEVLTLDNQLEVKAGAFGLTNVAV 366

Db 140 GTFHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDLNGLWNSLAVLPDAAF 190
QY 95 NDVLKLTIRLTKRNLKISRLPRAILDKMWLLQFLDHNALRDLDONLFOQLRNLOELGLN 154
Db 191 RGLSRLVLAGNRLAYLPALFSGLAELRELDLSRNALRAIKANVFVQLPRQLKLYLD 250
QY 155 QNQLSFLPANLFPSSRLKLLDLSRNLTHLPK-----GLLGAQVKLEKLLLYNQLTSTVD 210
Db 251 RNLIAAFAAGFAGLKAIRWLDLSHNRVAGLLEDTFPGLLGLRV-----LRLSHNAIASLR 306
QY 211 SGLLSNLCALTELRNLHRSVAPGADRLGNLSLTLGNNLESPPALFHVSVSR 270
Db 307 PRTEKDLHFLEELQGHNRIRQLAERSFEGLGQLEVLTLQHNQLOEVKAGAFGLTNAV 366
QY 271 LTFENPLEELPDVLFEGMAGRLWLNGLTHLSTLPAAAFRNLSGLQTLGLTRN----- 324
Db 367 MNLSCNCLRNLPQVFRGLGKLHSLHLEGSCIGRIRPHTFTGLSGLRRLFLKONGLVGIE 426
QY 325 -----PRLSALPRGVFGQLRELRLVLAHTNALAEALRDDLALRGLHURQ 367
Db 427 EQSLWGLAELELDLTSNQLTHLPHRLFQGLGKLEYLLLSNRNLAEPLADALPQORAFW 486
QY 368 VSLRHNRLRALPRTLFRNLSSL 389
Db 487 LDVSHNRLEALPNSLLAPLGR 508

RESULT 8

US-08-487-072A-49
; Sequence 49, Application US/08487072A
; Patent No. 6423684
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,072A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Insulin-like growth factor binding
; INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-487-072A-49

Query Match 21.6%; Score 419.5; DB 4; Length 605;
Best Local Similarity 33.0%; Pred. No. 3.7e-32;
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;
QY 38 GSAVH---IAELGLPTNLTHILFRMDQGLRNHFSGTMVLQRLMLSDSHISAIDPGTF 94
Db 140 GTFHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDLNGLWNSLAVLPDAAF 190
QY 95 NDVLKLTIRLTKRNLKISRLPRAILDKMWLLQFLDHNALRDLDONLFOQLRNLOELGLN 154
Db 191 RGLSRLVLAGNRLAYLPALFSGLAELRELDLSRNALRAIKANVFVQLPRQLKLYLD 250
QY 155 QNQLSFLPANLFPSSRLKLLDLSRNLTHLPK-----GLLGAQVKLEKLLLYNQLTSTVD 210
Db 251 RNLIAAFAAGFAGLKAIRWLDLSHNRVAGLLEDTFPGLLGLRV-----LRLSHNAIASLR 306
QY 211 SGLLSNLCALTELRNLHRSVAPGADRLGNLSLTLGNNLESPPALFHVSVSR 270
Db 307 PRTEKDLHFLEELQGHNRIRQLAERSFEGLGQLEVLTLQHNQLOEVKAGAFGLTNAV 366
QY 271 LTFENPLEELPDVLFEGMAGRLWLNGLTHLSTLPAAAFRNLSGLQTLGLTRN----- 324
Db 367 MNLSCNCLRNLPQVFRGLGKLHSLHLEGSCIGRIRPHTFTGLSGLRRLFLKONGLVGIE 426
QY 325 -----PRLSALPRGVFGQLRELRLVLAHTNALAEALRDDLALRGLHURQ 367
Db 427 EQSLWGLAELELDLTSNQLTHLPHRLFQGLGKLEYLLLSNRNLAEPLADALPQORAFW 486
QY 368 VSLRHNRLRALPRTLFRNLSSL 389
Db 487 LDVSHNRLEALPNSLLAPLGR 508
RESULT 9
US-09-538-092-1087
; Sequence 1087, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 1087
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: {0}...{0}
; OTHER INFORMATION: Polypeptide Accession Number P35858
US-09-538-092-1087

Query Match 21.6%; Score 419.5; DB 4; Length 605;
Best Local Similarity 33.0%; Pred. No. 3.7e-32;
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;
QY 38 GSAVH---IAELGLPTNLTHILFRMDQGLRNHFSGTMVLQRLMLSDSHISAIDPGTF 94
Db 140 GTFHTPALASGLSNN-----RLSRLEDGL-----FEGLSLWDLNGLWNSLAVLPDAAF 190
QY 95 NDVLKLTIRLTKRNLKISRLPRAILDKMWLLQFLDHNALRDLDONLFOQLRNLOELGLN 154
Db 191 RGLSRLVLAGNRLAYLPALFSGLAELRELDLSRNALRAIKANVFVQLPRQLKLYLD 250

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0950
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 603 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor bind.
INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33
US-08-190-802A-50

Query Match 19.7%; Score 382.5; DB 1; Length 603;
Best Local Similarity 28.7%; Pred. No. 1.5e-28;
Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;
QY 21 CPKTKCVVRD-----AAQCGGSAVAHIAELGPTNLTHTLLFRMDQGIILRNHSPGTMV 75
DB 41 CPVACTCSHDDYTDLSVFCSSKNLTHLPD-DIPVSTRALWLDGNNLSIPSAFQNLSS 99
QY 76 LQRLMLSDSHISAIDPGTFNDLVKLTIRLNKISRLPRAILDQVLLQELFDHNALR 135
DB 100 LDFNLQGSWLRSLEPQALLGQNLVYHLERNRLNLAVALGFTHTPSSLASLSLSSNLG 159
QY 136 DLQNLQFQRLNLOELGNLQNLSELPANLFSSIRELKLDLSRNNLTLPKGLGAQVK 195
DB 160 RLEEGLFQGLSHLWDLNLGWSLVLPDVTVFQGLGNLHVLVAGNKLTYLPALFCGLGE 219
QY 196 LEKLLYSNQLTSVDSGLLSNLGALTLELRNHLRSVAPGAF-----DRLG 242
DB 220 LRELDLSRNALRSVKANVFVHLPRQKLYLDRLNLTAVAPGAFGLGMKALRWLDLSHNRVA 279
QY 243 NLSS-----LTLSGNLLESPPALFLHVSSVSRLTLPENPLEEIPDVLFGEMAG 291
DB 280 GLMEDTFPGLGLHVLRLAHNAIASLRPTFKDLHFLBELQLGHNRIQIGERTFEGLGQ 339
QY 292 LRELWNGTHLSTLPAAAFRNLSGLQTLGTRPRLSALPRGVFGQLRELRVL----- 344
DB 340 LEVLTLDNQITEVRVGFAGSLFNVAVMNLSGN-CLRSLPRVFQGLDKLSLHLSHLSCL 398
QY 345 ---ALHT-----NALAEURD DALRGHLROVSRHNRRLRALPTLFRNLS 387
DB 399 GHVLTHTFAGLSGRLRLFRDNRNLSISSEQSLAGLSLLELDTNRLTHLPRLQFQGLG 458
QY 388 SL 389
DB 459 HL 460

RESULT 13
US-08-477-346-50
; Sequence 50, Application US/08477346

Patent No. 6262023
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSES: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 603 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor bind.
INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33
US-08-477-346-50

Query Match 19.7%; Score 382.5; DB 3; Length 603;
Best Local Similarity 28.7%; Pred. No. 1.5e-28;
Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;
QY 21 CPKTKCVVRD-----AAQCGGSAVAHIAELGPTNLTHTLLFRMDQGIILRNHSPGTMV 75
DB 41 CPVACTCSHDDYTDLSVFCSSKNLTHLPD-DIPVSTRALWLDGNNLSIPSAFQNLSS 99
QY 76 LQRLMLSDSHISAIDPGTFNDLVKLTIRLNKISRLPRAILDQVLLQELFDHNALR 135
DB 100 LDFNLQGSWLRSLEPQALLGQNLVYHLERNRLNLAVALGFTHTPSSLASLSLSSNLG 159
QY 136 DLQNLQFQRLNLOELGNLQNLSELPANLFSSIRELKLDLSRNNLTLPKGLGAQVK 195
DB 160 RLEEGLFQGLSHLWDLNLGWSLVLPDVTVFQGLGNLHVLVAGNKLTYLPALFCGLGE 219
QY 196 LEKLLYSNQLTSVDSGLLSNLGALTLELRNHLRSVAPGAF-----DRLG 242
DB 220 LRELDLSRNALRSVKANVFVHLPRQKLYLDRLNLTAVAPGAFGLGMKALRWLDLSHNRVA 279
QY 243 NLSS-----LTLSGNLLESPPALFLHVSSVSRLTLPENPLEEIPDVLFGEMAG 291
DB 280 GLMEDTFPGLGLHVLRLAHNAIASLRPTFKDLHFLBELQLGHNRIQIGERTFEGLGQ 339
QY 292 LRELWNGTHLSTLPAAAFRNLSGLQTLGTRPRLSALPRGVFGQLRELRVL----- 344
DB 340 LEVLTLDNQITEVRVGFAGSLFNVAVMNLSGN-CLRSLPRVFQGLDKLSLHLSHLSCL 398

QY	345	---	ALHT	-----NALAEIRDDALRGHLRQVSLRHNRLRALPRTLFRNL	387
Db	399	CHVRLHTFAGLSGLRRLFLRDNSSIEEQLAGLSELELDLTNRLTHLPQLFOGLG	458		
QY	388	SL	389		
Db	459	HL	460		
RESULT 14					
US-08-473-089-50					
; Sequence 50, Application US/08473089					
; Patent No. 6342368					
; GENERAL INFORMATION:					
; APPLICANT: Mochly-Rosen, Daria					
; APPLICANT: Ron, Dorit					
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses					
; TITLE OF INVENTION: Thereof					
; NUMBER OF SEQUENCES: 265					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Morrison & Foerster					
; STREET: 2000 Pennsylvania Avenue, NW					
; CITY: Washington					
; STATE: DC					
; COUNTRY: USA					
; ZIP: 20006-1812					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: Patent In Release #1.0, Version #1.25					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/473,089					
; FILING DATE: 07-JUN-1995					
; CLASSIFICATION: 435					
; ATTORNEY/AGENT INFORMATION:					
; NAME: MURASHIGE, KATE H.					
; REGISTRATION NUMBER: 29,959					
; REFERENCE/DOCKET NUMBER: 2550-0025.22					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (202) 887-1500					
; TELEFAX: (202) 887-0763					
; INFORMATION FOR SEQ ID NO: 50:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 603 amino acids					
; TYPE: amino acid					
; TOPOLOGY: unknown					
; MOLECULE TYPE: protein					
; HYPOTHETICAL: NO					
; ANTI-SENSE: NO					
; ORIGINAL SOURCE:					
; INDIVIDUAL ISOLATE: Insulin-like growth factor bind.					
; INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33					
US-08-473-089-50					
Query Match					
Best Local Similarity					
Matches 121; Conservative					
19.7%; Score 382.5; DB 3; Length 603;					
Pred. No. 1.5e-28;					
Mismatches 182; Indels 55; Gaps 7;					
QY	21	CPKTCCKVVRD	-----AAQCSGGSVAHIAELGLPTNLTHILLFRMDQGLRNHSPSGMTV	75	
Db	41	CPVACTCSHDDYDELSVFCSSKNLTHLPD-DIPVSTRALWDGNNLSIPSAAFQNLSS	99		
QY	76	LQRLMSDSHISAIDPTFNDLVKLTRLTRNKISRLPRAILDKNVLEQLFLDHNAIR	135		
Db	100	LDFLNQGSWLRSLEPQALLGLQNLXYLHLERNLENLAUGLFTHTPPSLASLSNLLG	159		
QY	136	DLQNLFFQQLRNQLQELGNQNSFLPANLFSRLRELKLLDLSRNNLTHLPKGLLGAQVK	195		
Db	160	RLEGLFQGLSHLDNLGNLSLVLPDVFQGLGNLHVLGAKNKLTYLQPALFQGLGE	219		
QY	196	LEXLLLYSNQUTSVDSGLLSNLGALTTELRLERNHLRSVAFGAP	-----DRUG	242	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:05:08 ; Search time 25 Seconds

(without alignments)
1497.133 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389

Perfect score: 1938

Sequence: 1 MLRSALLSAVLLIRAQFPF.....LRNRLRALPRTLFRNLSSL 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	76.2	560	2 A60164	platelet membrane
2	419.5	21.6	605	2 A41915	insulin-like growth
3	416.5	21.5	536	2 A34901	lysine carboxypept
4	413.5	21.3	605	2 JC5239	insulin-like growth
5	382.5	19.7	603	2 JC1282	insulin-like growth
6	382	19.7	603	2 JC6128	insulin-like growth
7	379.5	19.6	1531	2 T42218	slit-1 protein hom
8	367.5	19.0	907	2 JG0193	G protein-coupled
9	366	18.9	1523	2 TI3953	MEGF5 protein - ra
10	345	17.8	1389	2 TI3852	gene wheeler prote
11	339	17.5	907	2 JE0176	orphan G protein-c
12	337	17.4	1051	2 TI3174	gp150 protein - fr
13	337	17.4	1385	2 TI3887	tlr protein - frul
14	327	16.9	1031	2 A58322	glit cell membran
15	319.5	16.5	1489	2 B36665	slit protein 2 pre
16	319.5	16.5	1480	2 A36665	slit protein 1 pre
17	317.5	16.4	707	2 JC7763	neuronal leucine-r
18	317	16.4	361	2 A53860	chondroadherin pre
19	313	16.2	1066	2 TI5964	hypothetical prote
20	311.5	16.1	622	2 JC7973	synleucin - human
21	310	16.0	575	2 T29972	hypothetical prote
22	306.5	15.8	1134	1 A29944	chaoptin precursor
23	305.5	15.8	1097	2 A29943	Toll protein precu
24	305	15.7	981	2 T23395	hypothetical prote
25	302	15.6	610	2 T23836	hypothetical prote
26	301	15.5	594	2 T23841	hypothetical prote
27	300	15.5	789	2 T28714	hypothetical prote
28	300	15.5	1355	2 T28715	hypothetical prote
29	298.5	15.4	1119	2 AD1822	leucine-rich-repea

garp precursor - h
densin-180 - rat
hypothetical prote
hypothetical prote
hypothetical prote
disease resistance
hypothetical prote
protein AC7.2 [imp
leucine-rich alpha
Ras-binding protei
cell-surface molec
connecin precurs
gene flightless-1
disease resistance
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

A60164

platelet membrane glycoprotein V precursor - human

C:Species: Homo sapiens (man)

C:Date: 12-Jan-1993 #sequence revision 24-Feb-1994 #text change 09-Jul-2004

C:Accession: A48030; A60164; A35483; B35483; C35483; A60432; A47507; S34329

R:LANA, P.; Morales, M.; de La Salle, C.; Cazenave, J.P.; Clemetson, K.J.; Shimomura, T

J. Biol. Chem. 268, 20801-20807, 1993

A:Title: Cloning and characterization of the gene encoding the human platelet glycoprotein V: the

A:Reference number: A48030; MUID:94012616; PMID:8407908

A:Accession: A48030

A:Molecule type: DNA

A:Residues: 1-560 <LA2>

A:Cross-references: UNIPROT:P40197; EMBL:223091; NID:G912501; PIDN:CAA0637.1; PID:G9125

R:Shimomura, T.; Fujimura, K.; Maehama, S.; Takemoto, M.; Oda, K.; Fujimoto, T.; Oyama,

Blood 75, 2349-2356, 1990

A:Title: Rapid purification and characterization of human platelet glycoprotein V: the

A:Reference number: A60164; MUID:90275263; PMID:2350580

A:Accession: A60164

A:Molecule type: protein

A:Residues: 365-384, 'X', 386-390, 'X', 392-395, 'X', 397, 188-208, 'I', 210, 27-50, 'X', 52-53, 174-

, 'XX', 108, 'T', 61-72, 'TX', 75-77, 'V', 56-57, 'G', 478-487, 'X', 489-498, 'X', 500, 'X', 502-503, 'X',

R:Roeth, G.J.; Church, T.A.; McMullen, B.A.; Williams, S.A.

Biochem. Biophys. Res. Commun. 170, 153-161, 1990

A:Title: Human platelet glycoprotein V: a surface leucine-rich glycoprotein related to

A:Reference number: A35483; MUID:90321220; PMID:2372284

A:Accession: A35483

A:Molecule type: protein

A:Residues: 145-166, 'I', 168-169, 'X', 171-172 <ROT>

A:Note: this proteolytic fragment was designated peptide M392

A:Accession: B35483

A:Molecule type: protein

A:Residues: 121-129, 'W', 131-135, 466-468, 'X', 470 <RO2>

A:Note: this material was designated peptide M393 but may contain two peptides

A:Accession: C35483

A:Molecule type: protein

A:Residues: 252-266, 'H', 268-272, 'X', 274-279, 'I', 281-284, 'I', 286 <RO3>

A:Note: this proteolytic fragment was designated peptide M401

R:Zafar, R.S.; Walz, D.A.

Thromb. Res. 53, 31-44, 1989

A:Title: Platelet membrane glycoprotein V: characterization of the thrombin-sensitive g

A:Reference number: A60432; MUID:89162331; PMID:2922700

A:Accession: A60432

A:Molecule type: protein

A:Residues: 477-478, 'FX', 481-485, 'E', 487, 'V', 489-492, 'NQ', 495, 'E', 497-498 <ZAF>

R:Hickey, M.J.; Hagen, F.S.; Yagi, M.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 90, 8327-8331, 1993

A:Title: Human platelet glycoprotein V: characterization of the polypeptide and the rel.

A:Reference number: A47507; MUID:93391348; PMID:7690959

A:Accession: A47507

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A;Residues: 1-560 <RES>
A;Cross-references: GB:111238; NID:G388759; PIDN:AAA03069.1; PID:G388760
C;Comment: This platelet membrane protein is a substrate for thrombin.
C;Comment: The amino end of the intact protein is blocked.
C;Comment: This protein is absent in Bernard-Soulier syndrome.
C;Genetics:
A;Gene: GDB:G95
A;Cross-references: GDB:230236; OMIM:173511
A;Map position: Spiter-5qter
C;Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane protein

Query Match 76.2%; Score 1477; DB 2; Length 560;
Best Local Similarity 76.1%; Pred. No. 1-7e-94;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

Qy	1	MLRSALLSVALILRAQPPCTKCVVDRAQCGSGVAHIEGLPTNLTTHILLFRM	60
Db	1	MLRGTLCAVLGLRAQPPCTKCVVDRAQCGSGVAHIEGLPTNLTTHILLFRM	60
Qy	61	DOGLRNHPSFGMTVLQRLMLSDSHISAIIDPGTNDLVKLTIRLNKIKRLPRAILDK	120
Db	61	GRGVLOSQSGMTVLQRLMLSDSHISAIIDPGTNDLVKLTIRLNKIKRLPRAILDK	120
Qy	121	MVLLEQLFQHNALRDQNLQOLNQLGNGQSLFIPANLFPSSIRELKLDLSRN	180
Db	121	MVLLEQLFQHNALRDQNLQOLNQLGNGQSLFIPANLFPSSIRELKLDLSRN	180
Qy	181	NLTHLPKGLGAGVAKLERLLHNSRLVSLDGLNSLGLTALTELOPHRNHISIAFGADR	240
Db	181	NLTHLPKGLGAGVAKLERLLHNSRLVSLDGLNSLGLTALTELOPHRNHISIAFGADR	240
Qy	241	LGNISSUTLSCNLESIPPAFLPHVSVSRILTFENPLEELPDVLFGEAGRLINLNGT	300
Db	241	LGNISSUTLSCNLESIPPAFLPHVSVSRILTFENPLEELPDVLFGEAGRLINLNGT	300
Qy	301	HLSTLPAAPRNLSGLTGLTNPRLSALPRGVFGQRLRVLALHTNALAEIRDALR	360
Db	301	HLSTLPAAPRNLSGLTGLTNPRLSALPRGVFGQRLRVLALHTNALAEIRDALR	360
Qy	361	GLGHLRQVSLRHNRLALPRTLPRLNSSL	389
Db	361	GLGHLRQVSLRHNRLALPRTLPRLNSSL	389

RESULT 2
A41915
Insulin-like growth factor-binding complex acid-labile chain precursor - human
A;Alternate names: Acid-labile Subunit (ALS)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A41915
R;Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.
Mol. Endocrinol. 6, 870-876, 1992
A;Title: Structure and functional expression of the acid-labile subunit of the insulin-like growth factor-binding complex
A;Reference number: A41915; MUID:92357025; PMID:1379671
A;Accession: A41915
A;Status: Preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-605 <LEO>
A;Cross-references: UNIPROT:P35858; GB:M86826; NID:gi84807; PIDN:AAA36047.1; PID:gi84808
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBI:110171)
F;75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;99-122/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F;315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F;339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F;363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F;387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F;411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
F;435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
F;459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>
F;483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>
F;507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>

Query Match 21.6%; Score 419.5; DB 2; Length 605;
Best Local Similarity 33.0%; Pred. No. 2.1e-21;
Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

Qy	38	GSVAH---IAELGPTNLTTHILLFRMDGTLRNHPSFGMTVLQRLMLSDSHISAIIDPGTF	94
Db	140	GFATPTPALASGLSNN-----RLSRLEDGL-----FEGLSLMDLNLGNLSLAVIPDAF	190
Qy	95	NDLVKLTIRLNKIKRLPRAILDKMVLLEQLFQHNALRDQNLFOQLRNQLGNGT	154
Db	191	RGLSRLVLAGNRLAYLQPALFSGLAELRELDLSGNALRAIKANVFQLPRLQKLYLD	250
Qy	155	QNQLSFLFANLFPSSIRELKLDLSRNLTLPK---GLGAGVAKLERLLHNSRLVSLDGLNSLGLTALTELOPHRNHISIAFGADR	210
Db	251	RNLIAAVAPAGFLGLKALRWLDSHNRVAGLEDTFFELGLLRV---LRLSHNATASLR	306
Qy	211	SGLLSNLQALTELRNLERNHLSVAPGAPDRNLGNLSSLTSGNLESPLPALFHVSVSR	270
Db	307	PRTFKDLHFLBELQGLHNRIRQLAERSFEGLGQLEVLTLQHNQLEQVYKAGFLGTNVAV	366
Qy	271	LTLFENPLEELPDVLFGEAGRLINLNGTGLSTLPAAAFRNLSGLTGLTN-----	324
Db	367	NMLSGNCLRNLPQVFGGLGKLSHLEGSCIGIRPHFTGSGRLFLKDNGLVGLIE	426
Qy	325	-----PRLSALPRGVFGQRLRVLALHTNALAEIRDALRGLGHLRQ	367
Db	427	EQSLMGLAELELDLTSNQLTHPLRFLFOGLGKLEVLTLQHNQLEQVYKAGFLGTNVAV	486
Qy	368	VSLRHNRLALPRTLPRLNSSL	389
Db	487	LDVSHNRLALPNSLLAPLGLRL	508

RESULT 3
A34901
lysine carboxypeptidase (EC 3.4.17.3) 83K chain - human
C;Species: Homo sapiens (man)
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 09-Jul-2004
C;Accession: A34901
R;Tan, F.; Weerasinghe, D.K.; Skidgel, R.A.; Tamei, H.; Kaul, R.K.; Roninson, I.B.; Schif
J. Biol. Chem. 265, 13-19, 1990
A;Title: The deduced protein sequence of the human carboxypeptidase N high molecular weight
A;Reference number: A34901; MUID:90094386; PMID:2378615
A;Accession: A34901
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-536 <TAN>
A;Cross-references: UNIPROT:P22792; GB:J05158; NID:gi79935; PIDN:AAA51921.1; PID:gi79936
A;Gene: GDB:ACBP
A;Cross-references: GDB:127893
A;Map position: 6q25.3-6q26
C;Keywords: hydrolase; metallo-carboxypeptidase
F;101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;221-244/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;245-268/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;269-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;317-340/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:341-364/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>

Query Match 21.5%; Score 416.5; DB 2; Length 536;
Best Local Similarity 32.3%; Pred. No. 2.9e-21;
Matches 115; Conservative 53; Mismatches 163; Indels 25; Gaps 3;
QY 22 PKTKCVVRDAQCGSGSVAH-----AELGFTWLTHTLF 58
Db 2 PMGDCVQEVF-CSDDELATVPLDIPPYTKNIIFVETSTLTETRAFGSNPLTKVFL 60
QY 59 RMDQILRNHSFGSMVTLQRLMLSDSHISAIDPGTFNDLVKLTILTRNKISRLPRAIL 118
Db 61 DTQLCQRPDAFGGLPLEDLEVTGSSFLNLTWNISLTSKLTILFNMLPEGF 120
QY 119 DKWVLLBQLFDHNALEDLQNLQQLRNQELGNQNLQSLFANLFSLSRELKLLDLS 178
Db 121 QHLAALESHLQGNQALPRLFPQPLTHLKLNLNAQNLLAQLPEELFHLPLTSLQTLKS 180
QY 179 RNNLTPLKGLGAQVKLEKLLYSNQLTSVDSGLSLNGALTELRLRNHLRSVAPGAF 238
Db 181 NNALSGLPGQVFGKLSQLSELDSDNNISELPQVFSOLFCLRLWLQRNAITHLPLSTF 240
QY 239 DRUGNLSLTLSGNLLSPLPALFLHVSSVSRLTLFENPLEELPDVLFEGMAGRLWLNL 298
Db 241 ASLGNLFLSLQWMLRVLPAGLPFAHTPCLVGLSLTHNQLETVTEGTFHLSNLSMLLS 300
QY 299 GTHLSTLTPAAAFNLGLTGLTRNPRLSALPRGVQGLRELRLVALHTNALAEL 354
Db 301 YNAITHLPAGIFRDLLELVKYLGSN-NLTALHPALFQNLKLELLSKNQTLTL 355

RESULT 4

JC5239
insulin-like growth factor acid-labile chain - baboon
C:Species: Pabo sp. (baboon)
C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
C:Accession: JC5239
R:Delhanty, P.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 227, 897-902, 1996
A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like
A:Reference number: JC5239; MUID:97040714; PMID:8886027
A:Contents: liver
A:Accession: JC5239
A:Molecule type: mRNA
A:Residues: 1-605
C:Comment: This factor is structurally related to proinsulin and have insuline-like meta

Query Match 21.3%; Score 413.5; DB 2; Length 605;
Best Local Similarity 29.8%; Pred. No. 5.4e-21;
Matches 141; Conservative 56; Mismatches 167; Indels 109; Gaps 8;
QY 21 CPKTKCVVRDAAQ-----CSG----- 37
Db 41 CPATCASYDDEVNELSVFCSSRNLTLPDGPICGTQALWLDNSNLSSTPPAFLNLSL 100
QY 38 -----GSAVHAIEGLPNLTHILLFRMDQILRNHGFSGMTVLQRLMLSDSHIS 87
Db 101 AFLNLQGGQLGSLPEQALLGL-ENLCHLHLERNQLRSVAVGTFAYPALALLGLSNRLS 159
QY 88 AIDPGTFNDLVKLTILTRNKISRLPRAI-----LDKXVL----- 123
Db 160 RUDGGLFEGLGNLWDLNLGNLSLAVLPDAAFRGLGGLRELVLGNRLAYLQPALFSGLA 219
QY 124 LEQFLFDHNALEDLQNLQQLRNQELGNQNLQSLFANLFSLSRELKLLDLSRNNLT 183
Db 220 LRELDLSRNALRAIKANVPAQLPRLQKLYLDRLNLIAAVAPGAFGLKALRWLDLSHNRVA 279
QY 184 HLPK----GLLGAQVKLEKLLYSNQLTSVDSGLSLNGALTELRLRNHLRSVAPGAF 239
Db 280 GLIEDTFPGLGLGV-----LRUSHNAIASLRPTPDHLFELQGHNRIRQLAESPE 335
QY 240 RLGNLSLTLSGNLLSPLPALFLHVSSVSRLTLFENPLEELPDVLFEGMAGRLWLNL 299

Db 336 GLGQLEVLTDHNLQEVKVGAFGLGTNNVAVMNLSGNCLRNLPQVPRGLKGLSHLHLEG 395
QY 300 THLSTLTPAAAFNLGLTGLTRN-----PRLSALPRGVQ 336
Db 396 SCIGIRPHHTFAGTSGRLFLKONGLVGIEQSLWGLAELELDLTNSQTHLPHQLFQ 455
QY 337 GRLRELRLVALHTNALAELRDDALRGHLRQVSLRHNELRALPRTLFRNLSL 389
Db 456 GLGKLEYLLSHNLRAELPADALGPLQAFWLVDVSHNEALPGSLLASLGR 508

RESULT 5

JC1282
insulin-like growth factor-binding protein acid labile chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: JC1282
R:Dai, J.; Baxter, R.C.
Biochem. Biophys. Res. Commun. 188, 304-309, 1992
A:Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac
A:Reference number: JC1282; MUID:93038676; PMID:1384485
A:Accession: JC1282
A:Molecule type: mRNA
A:Residues: 1-603
A:Cross-references: UNIPROT:P35859; GB:S46785; MID:g258002; PIDN:AAE23770.2; PID:g570593
A:Experimental source: liver
A:Note: The authors translated the codon AAG for residue 63 as Arg, AAA for residue 205
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-603/Product: insulin-like growth factor binding protein, acid labile chain #status
F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 19.7%; Score 382.5; DB 2; Length 603;
Best Local Similarity 28.7%; Pred. No. 7.5e-19;
Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;
QY 21 CPKTKCVVRD-----AAQCSGSGVAHAEGLTNTLTHILLFRMDQILRNHGFSGMTV 75
Db 41 CPVACTCSHDDYTDLSVFCSSKNLTHLPD-DIPVSTALWLDGNNLSISIPSAFQNLSS 99
QY 76 LQRLMLSDSHISAIDPGTFNDLVKLTILTRNKISRLPRAILDKMVLLEOLFHDHNA 135
Db 100 LDFLNLQGSWLRSLEPQALLGLQNLVYLHLERNLRLNLAVLGFTHTPSLASLSLSSNLLG 159
QY 136 DLDQNLFOQLRNQELGNQNLQSLFANLFSLSRELKLLDLSRNNLTPLPKGLLGAQVK 195
Db 160 RLEELFGCLGSHLWDLNLGNLSLVLPDTPVQGLNHLVLAGNKUITYLQPALFCGLGE 219
QY 196 LEKLLYSNQLTSVDSGLSLNGALTELRLRNHLRSVAPGAF-----DRLG 242
Db 220 LRELDLSRNALRSVKANVFVHLPRQLKLYLDRLNLTITAVAPGAFGLGMKALRWLDLSHNRVA 279
QY 243 NLSS-----LTLISGNLLESPLPALFLHVSSVSRLTLFENPLEELPDVLFEGMAG 291
Db 280 GLMEDTFPGLGLHVLRLAHNAIASLRPTFKDHLFEEQLQGHNRIRQLGERTFEGLGQ 339
QY 292 LRELWNLTHLSTLTPAAAFNLGLTGLTRNPRLSALPRGVQGLRELRLV----- 344
Db 340 LEVLTNDNQITEVRVAGFSGLFENVAVMNLSGN-CLRSLPERVFGQLKGLSHLHLSCL 398
QY 345 ---ALHT-----NALAEIRDDALRGHLRQVSLRHNELRALPRTLFRNLS 387
Db 399 GHVRLHTFAGLSGLRRLFLRDNSISSIEOSLAGLSLELDLTNTNLTPLPQLFQGLG 458
QY 388 SL 389
Db 459 HL 460

RESULT 6

JC6128
insulin-like growth factor binding complex acid labile chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004

Db 337 -GAKISSLPQAVCDQLNQLVLDLSYNLLEDL--PSLSGGCKQKQIDLRNNEIYEIKGST 393
Qy 383 FRNLSSL 389
| : :
Db 394 PQQLFNL 400

RESULT 9
TI3953
MEGF5 protein - rat
N:Alternate names: slit protein homolog
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T13953
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
E:Genomics 54, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: Z14126; MUID:98360089; PMID:9693030
A:Accession: T13953
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1523 <NA>
A:Cross-references: UNIPROT:O88280; EMBL:AB011531; NID:G3449291; PIDN:BAA32461.1; PID:G3449291
C:Genetics:
A:Gene: MEGF5
C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein x

Query Match 18.9%; Score 366; DB 2; Length 1523;
Best Local Similarity 20.2%; Pred. No. 3.4e-17;
Matches 120; Conservative 85; Mismatches 180; Indels 210; Gaps 8;

Qy 3 RSALLSVALLRAPQ-FPCPKTKCVVRDAACSGGSAHIAELGLPTNLTLLIFRMD 61
| : : : : :
Db 15 RLALALASILSGPPAACTKCTCSAA-SVDCHGLGLRAVPR-GIPNAERLDDRN 72
| : : : : :
Qy 62 QGILRNHSFGMTVQLRLMLSDSHISAIDPGTNDLVKLTIRLNKISRPLAILDKM 121
| : : : : :
Db 73 ITRITKMDFTGLKNRLVHLHEDQVSVIERGAFQDLKQLERLNKLNKLQVLPQLFQST 132
| : : : : :
Qy 122 VLEQLFLDNRALDLDONLFOOLRNLOELGNQNL-SFLPANLPSLSRELKLLDSRN 181
| : : : : :
Db 133 PKLTRLDLSENOIGTPKRAFRGVTGVKNLQNNHISCIEDGAFRALDLEILLUNNN 192
| : : : : :
Qy 182 LTHLPKGLLGAQVKLEKLLYSNOL-----TSVDSGLLSNLG----- 206
| : : : : :
Db 193 ISRLVTSFNHMPKIRTLRLSHNHLVCDCHLAWLSDWLQRRTIGQFTLCMAPVHLRGFS 252
| : : : : :
Qy 207 -----TSVDSGLLSNLG-----ALTEL 223
| : : : : :
Db 253 VADVQKKEYVCPGPHSEAPACNANSICPSACSNIVDCRGKGLTEIPANLPEGIVEI 312
| : : : : :
Qy 224 RLERNHLRSVAPGAFDRGLNLSLTLSGNLLESPLPALFLHVSSVSRLTLFENPLEELPD 283
| : : : : :
Db 313 RLEQNSIKSIPAGAFICYKKLRIDISKQISDIADPAQGLKSLTSLVYGNKITEIPK 372
| : : : : :
Qy 284 VLGEMAGRELWNGTHSLTPAAAFRLNSGL-----QTL 319
| : : : : :
Db 373 GLFDGLVSLQLLLNANKINCLEWNTFDQLONLNLSDYDNKLQTSKGLFAPLQSIQTL 432
| : : : : :
Qy 320 GLTRNP----- 325
| : : : : :
Db 433 HLAQNPFVCDCHLKLWADVLQDNPITSGARCSPRLANKRISQIKSKKFKCSCSEDYR 492
| : : : : :
Qy 326 -----RLSALP-----RGV 334
| : : : : :
Db 493 NRFSECFMDLVCFEKCRCCEGTIVDGSNQLSRIPSHLPEYTTDLRLNDNDIAVLEATGI 552
| : : : : :
Qy 335 FQGLRELRLVLAHTNALAELRDDALRGLHQLQVSLRHLNRALPRTLFRNLSSL 389
| : : : : :
Db 553 FKLLPNLRKINLSNNRIKEVREGAFDGAAGVQELMTGNQLTETMGRMFRGUSGL 607
| : : : : :
RESULT 10

TI3852

gene wheeler protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T13852

R:Eldon, E.; Kooyer, S.; D'evilyn, D.; Duman, M.; Lawinger, P.; Botas, J.; Bellen, H.
E:Development 120, 885-899, 1994

A:Title: The Drosophila 18 wheeler is required for morphogenesis and has striking simila

A:Reference number: Z17796; MUID:95324375; PMID:7600965

A:Accession: T13852

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1389 <EUD>

A:Cross-references: UNIPROT:Q24591; EMBL:L23171; NID:G415682; PID:G1019104; PIDN:AAA7920

C:Genetics:

A:Gene: wheeler

A:Cross-references: FlyBase:FBgn0004364

Query Match 17.8%; Score 345; DB 2; Length 1389;
Best Local Similarity 30.4%; Pred. No. 8.6e-16;
Matches 113; Conservative 57; Mismatches 150; Indels 52; Gaps 8;Qy 65 LRNHSFGMTVQLRLMLSDSHISAIDPGTNDLVKLTIRLNKI-----KVVLEQLFLDHNALRDLDQNL 110
| : : : : :
Db 136 LHQSFQGLKELSELHGLGNNIRQLPEGVCWCMPSLQLLNTQNRISAEFLGFSEKLC 195
| : : : : :
Qy 111 -SRLPRA-----ILD-----KVVLEQLFLDHNALRDLDQNL 142
| : : : : :
Db 196 GSALSNANGAVSGGSELQTLDSVFNELSLPDWAGASRLRELQTLTSLQHNISTLAPNAL 255
| : : : : :
Qy 143 QOLRNLOELGNQNL-SFLPANLPSLSRELKLLDLSRNLTPLKGLLGAQVKLEKLL-- 200
| : : : : :
Db 256 AGLSLRLVNTSYNHVSLPSEAAGKELRELHQLQNDLYELPKGLLH---RLQQLLV 312
| : : : : :
Qy 201 -LYSNQLTLS--VDSGLLSNLGALTTELRLNHLNRSVAPGAFDRLGNLSLTLSGNLLES 257
| : : : : :
Db 313 DLGSLQTLSHVDNSTFAGLRILVNLNLSNALTTRIGSKTFKELYVFLQILDMRNSIGHI 372
| : : : : :
Qy 258 PPALFLHVSSVSRLTLFENPLEELPDVLFEMAGRELWNGTHSLTPAAAFRLNSGLQ 317
| : : : : :
Db 373 BEGAFPLYNLHTLNLAEENRLHTLDNRIFNGLYVLTLLNNNLVSIVESQAFRCNCDLK 432
| : : : : :
Qy 318 TLGLTRNPLSALPRGVFQGLRELRLVLAHTNALAELRDDALRGLHQLRQVSLRHLNR 377
| : : : : :
Db 433 ELDLSSN-QLTEVPEAV-QDLNMLKTLDDGENQISEFKNNTFRNLNQLTGLRLDNRIGN 490
| : : : : :
Qy 378 LPRILFRNLSSL 389
| : : : : :
Db 491 ITVGMFQDLPL 502
| : : : : :
RESULT 11

JEO176

orphan G protein-coupled receptor precursor - human

C:Species: Homo sapiens (man)

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004

C:Accession: JEO176

R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

E:Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A:Title: Identification and cloning of an orphan G protein-coupled receptor of the glyco

A:Reference number: JEO176; MUID:98308104; PMID:9642114

A:Accession: JEO176

A:Molecule type: mRNA

A:Residues: 1-907 <MCD>

A:Cross-references: UNIPROT:O75473; GB:AF062006; NID:G3366801; PIDN:AAC28019.1; PID:G336

C:Genetics:

A:Gene: HG38

A:Map position: 12q22-23

F:1-21/Domain: signal sequence #status predicted <SIG>

F:562-583/Domain: transmembrane #status predicted <TM1>

F:594-616/Domain: transmembrane #status predicted <TM2>

F:639-660/Domain: transmembrane #status predicted <TM3>

F:601-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 17.5%; Score 339; DB 2; Length 907;
Best Local Similarity 27.7%; Pred. No. 1.3e-15;
Matches 113; Conservative 59; Mismatches 158; Indels 68; Gaps 10;

QY 13 LLRAQPPPKTKCKVVRDAAQCGGSAHI--ABIG---LPTNTHILLPRMDCGLRN 67
DB 30 LLRG-----CPETHCHC-----EPDGMRLRVDCSDGLSELPSNLVFTSY----- 70
QY 68 HSPSGMTVLQRLMLSDSH:SAIDPGTFNDVLKTLRLTRNKRIPRAILDKMLLEQL 127
DB 71 -----LDSMNNISQLLPNPLPSLRLEELRLAGNALTYPKAGATGLYSKVL 119
QY 128 FLHNALRDQNLQFQNLNQLBELGQNLQFNLFLPANLFSSIRELKLDDLRNNLTHLPK 187
DB 120 MLQNNQLRHVPTALQNLRLSLQSLDANHI SYVPPSCFGLSLRLHMLDNLALTEIPV 179
QY 188 GLLGAQVLEKLLLYGNQLTSVDSGLSLNLGALTLELRNHLRSVAPCAFDRNLSSL 247
DB 180 QAFRLSALQAMTALNKHHPDYAFGNLSSLVVHLNHNRIHSLGKCKDGLHSLETL 239
QY 248 TLGNLLESUPPAL-----FLHVSVSRLTLFENPEELPDV 284
DB 240 DLYNNLDSEFTTAIRTLNKLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNP:QFVGRS 299
QY 285 LFGEMAGRELWNGHGLTLPAAAFRLNSG---LQTLGTRNPRLSALPRGVFQGLREL 341
DB 300 AFQHLPELRLTLNGASQIT-----EPDLTGTALESLETLT-GAGISLRLTVCNQPLNL 354
QY 342 RVLAHTNALAEURDALLGLGLHQLROVSURHNRALPRTLFRNLSSL 389
DB 355 QVLDLSYNLLELDPSESV--CQKLOKIDLRHNEIYEIKVDTFOQLLSL 400

RESULT 12.

Tl3174
gpl50 protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: Tl3174
R:Tian, S.S.; Zinn, K.
J. Biol. Chem. 269, 28478-28486, 1994
A:Title: An adhesion molecule-like protein that interacts with and is a substrate for a
A:Reference number: Z17630; MUID:95050638; PMID:7961789
A:Accession: Tl3174
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1051 <TIA>
A:Cross-references: UNIPROT:Q24007; EMBL:U15220; NID:G595859; PID:G595860; PIDN:AAA61798
A:Experimental source: strain Oregon R
C:Genetics:
A:Cross-references: FlyBase:FBgn0013272
A:Map position: 2

Query Match 17.4%; Score 337; DB 2; Length 1051;
Best Local Similarity 27.1%; Pred. No. 2.1e-15;
Matches 117; Conservative 65; Mismatches 189; Indels 60; Gaps 10;

QY 16 AQPFPCKTKCKVVRD-----AAQCS-----GGSAHIA----- 44
DB 228 SEDVPCRFCCQ-ARNVSYLVATSLRDMGIXQKSGSDITLDVTNVCVKYPIILMGNNF 286
QY 45 -ELGLPTNLTHILLFRMDQGLIRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLT 103
DB 287 QNLGL-KNVASIKIANCTLEYLHAEAFHGLNELYAVNLTDVGLAIINPTFVGNKLRML 345
QY 104 RLTRNKISRUPRA-ILDKMVLLEQLFDHNLARDLQNLFOQLRNLQELGNQNLGFLP 162
DB 346 TISGNDLSVMSSTHLLKSSSIEELDFENNLMLNPKAFSHLSNVVINSQSLKKLP 405

QY 163 ANLFSSIRELKLDDLRNNLTHLPKGLLGAQV-----KLEKLLLY 202
DB 406 EKAFKVTLLBELDLSYNSLTELPRDFNGTTLSILHKYKNTNGDLHFGTKDQLQDLUS 465
QY 203 SNQLTSVDSGLSLNLGALTLELRNHLRSVAPCAFDRNLSSLTSLGNLLESPLPALF 262
DB 466 FNSIVQVHSMFDMKPGGLTNLKGNGIKKIQDPSFLTKNLARHIDLSINDLDQISGMFL 525
QY 263 LHVSVSRLTLFENP-LEELPDVLFEGMAG---LRELWNGHGLTLPAAAFRLNSGLT 318
DB 526 FKNSELDVIRLNDPRLSQLPTDGLFSYSGEFTVYVYLDISNCAIGLGHKAFSTWPHLT 585
QY 319 LGLTENPRLSALPRGVFQGLRELRLVALHTNALAEFLDRLALRGHLRLQVSRHNRAL 378
DB 586 LKLANW-NINHLPREITGLRKLDDLSNNLITRMDLLIFMDNGELTKLSLGNPISRL 644
QY 379 PRTLFRNLSSL 389
DB 645 SVRLFPLHLQ 655

RESULT 13

Tl3887
tlz protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: Tl3887
R:Chiang, C.; Beachy, P.A.
Mech. Dev. 47, 225-239, 1994
A:Title: Expression of a novel Toll-like gene spans the parasegment boundary and contrib
A:Reference number: Z17805; MUID:95151581; PMID:7848870
A:Accession: Tl3887
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1385 <CHI>
A:Cross-references: UNIPROT:Q24591; EMBL:S76155; NID:G913247; PID:G913249; PIDN:AA833383
C:Genetics:
A:Cross-references: FlyBase:FBgn0004364
A:Note: tlr

Query Match 17.4%; Score 337; DB 2; Length 1385;
Best Local Similarity 29.8%; Pred. No. 3.1e-15;
Matches 111; Conservative 57; Mismatches 152; Indels 52; Gaps 8;

QY 65 LRNHSFGMTVLQRLMLSDSHISAIDPGTFNDLVKLTLELRNKR----- 110
DB 136 LHGQSFQGLKELSELHGLDNNIRQLPQGVWCMPSLQLLNTQNRIRSAEFLGFSEKLCA 195
QY 111 -SRLPRA-----ILD-----KVVLEQLFDHNLARDLQNL 142
DB 196 GSALSNANGAVSGSGEQLTLDVSNELRSLPDWAGASRLRLQLTSLQHNNTSLTAPNAL 255
QY 143 QQLRNLQELGNQNLQFLPANLFSSIRELKLDDLRNNLTHLPKGLGQAVKLEKL-- 200
DB 256 ACLSLRLVNTSYNNHLYSLPSEAFAGNKLRELHQLQNDLYELPKGLH---RLEQLLVL 312
QY 201 -LYSNQLTS--VDSGLSLNLGALTLELRNHLRSVAPCAFDRNLSSLTSLGNLLES 257
DB 313 DLGSGNQLTDHVDNSTFAGIRLVNLNNAITRGSKTFKELYFLQIDMRNNSIGHI 372
QY 258 PPALFLHVSSVSRLTLFENPLEELPDVLFEGMAGRELWNGHGLTLPAAAFRLNSGLQ 317
DB 373 EBGAFPLYNLHTLNAENRLHLDNRIFNGLYVLTGLTNLNLVSVESQAFRNCSDLK 432
QY 318 TLGLTRNPRLSALPRGVFQGLRELRLVALHTNALAEFLDRLALRGHLRLQVSRHNRAL 377
DB 433 ELDLSSN-QLTEVEPAA-QDLSMLKTLDELGENQISEFKNNTFRNLNQLTGLRLIDNRGN 490
QY 378 LPRTLFRNLSSL 389
DB 491 ITVGMFQDLPRL 502

RESULT 14

A36532

Glia cell membrane glycoprotein LIG-1 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 09-Jul-2004

C:Accession: A36532

R: Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.

J. Biol. Chem. 271, 22522-22527, 1996

A:Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in

A:Reference number: A36532; MUID:96394313; PMID:8798419

A:Accession: A36532

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1091 <SUZ>

A:Cross-references: UNIPROT:P70193; GB:D78572; MUID:91545806; PID:BA11416.1; PID:G15458

F:36-61/Domain: proteoglycan amino-terminal homology <PAH>

F:71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F:262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F:286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F:334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F:358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F:385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F:409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F:440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match

16.9%; Score 327; DB 2; Length 1091;

Best Local Similarity 28.4%; Pred. No. 1.1e-14;

Matches 114; Conservative 78; Mismatches 182; Indels 28; Gaps 11;

Qy 6 LLSAVLALL-----RAQP-FPCPKTKCVWRDAAQCSGGVYAHIAELGLPT-----N 51

Db 19 LLWLLLLLPESAGAQAPRAPCAAACTC-AGNSLDCSGRGLATLPR-DLFSWTRSLN 76

Qy 52 LTHILLFRMOGILRNHSFGMTVLQRLMSDHSIHSAIDGTNDVLKLTILTRNKIS 111

Db 77 LSYNRLSEIDSA-----AFEDLTNLQVYLSNELTAT-PSLGTASIGVVSFLQHNKIL 130

Qy 112 RLPRAILDKMVLLEQLFDHNAIRDLDQNLFQQLRNQLGILNQNLSPANLFSLL-R 170

Db 131 SVDSGQSKVSLSEVLDSNNITEIRSSCFPNGLRIRELNLASNIISLESAPGLSR 190

Qy 171 ELKLLDLRNHLTHPLKGLGAQVKLEKLLLYNQNLTSVDSGLLSNLGALTLELRNHL 230

Db 191 SLTLRLSKNRITQLPVKAPLP-RLTQLDLNENRIRIEGLTFQGLDLSLEVLRLQNNI 249

Qy 231 RSVAPGAFDRIGNLSSLTLSGNLLESPPALFLHVSVSRLTLFENPLEELPDLVFGMA 290

Db 250 SRLTDGAFGLSKMHLVLEYSLSVENSGLGLTALHQLHLSNNSIRIQRDGSPCQ 309

Qy 291 GLRELWNGTHLSTLPAAAFRNLSGLTGLTRNPLRSALPRGVFGQLRELRLVLAHTNA 350

Db 310 KLHELILSFNNLTLDEESLAELSSILRLSHN-AISHIAEGAFKGLSLRVLDLDHNE 368

Qy 351 LAELRDD--ALRLGHLROVSLRHRNRLRALPTLFRNLSSL 389

Db 369 ISGTIETSGAFTGLDNLKLTIFGNKIKSVAKRAFSGLES 410

RESULT 15

B36665

slit protein 2 precursor - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002

C:Accession: B36665

R: Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Teakonas, S.
Genes Dev. 4, 2169-2187, 1990

A:Title: slit: an extracellular protein necessary for development of midline glia and c

A:Reference number: A36665; MUID:91099665; PMID:2176636

A:Accession: B36665

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1469 <ROT>

A:Cross-references: GB:X53959

C:Genetics:

A:Gene: FlyBase:slit

A:Cross-references: FlyBase:FBgn0003425

C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>

F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>

F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>

F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>

F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>

F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>

F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>

F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>

F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>

F:1028-1061/Domain: EGF homology <EGF>

F:1068-1099/Domain: EGF homology <EGF2>

F:1115-1148/Domain: EGF homology <EGF1>

Query Match

16.5%; Score 319.5; DB 2; Length 1469;

Best Local Similarity 29.3%; Pred. No. 5.4e-14;

Matches 118; Conservative 64; Mismatches 156; Indels 65; Gaps 14;

Qy 21 CPKTKCVWRDAAQCSGGSV-----AHIAELGLP-TNLTHILLFRMOGILRNHSFS 71

Db 73 CPVCSCTGLN-VDCSHRGLTSVPRKISADVERLELQGNLIT-----VIYETDFQ 121

Qy 72 GMTVLQRLMSDHSISAIIDPGTFNDVLKLTILTRNKISRLPRAILDKMWLLQLFLDH 131

Db 122 RLTKLRMLQLTNQIHT-ERNSFQDLVSLERLIDISNNVITTVGRRVFKGAQSLRLQLDN 181

Qy 132 NALRDLQNLFCQLENLQELGILNQNLSPANLFSLSRELKLDLSN-----NLTHL 185

Db 182 NOITCLDSEAFKGLVELEILTNNNNLTSLPHNIFGGLRLRLRLSDNPPFACDCHLSWL 241

Qy 186 PKGLIGAQVKLEKLLYS-----NQLTSVD-----SGLLSNLGALTLELRN- 228

Db 242 SRFLESA---TRLAPYTRCQSPQLKQNVADLHDQEFKCSGLTEH--APMECAENSC 295

Qy 229 -HLRSVAPGAFD-RLGNLSLTLGSLNLESILPPALFLHVSVSRLTLFENPLEELPDLVF 286

Db 296 FPHCRCADGIVDCREKSLTSVPV-----TLF-----DDTTDVRLQNFITELPKPSF 342

Qy 287 GEMAGRELWNGTHLSTLPAAAFRNLSGLTGLTRNPLRSALPRGVFGQLRELRLVAL 346

Db 343 SSFRLRLRIDLSNNISRIADHLSGKQLTTLVLYGN-KINKDPSGVFKGLSLRLULL 401

Qy 347 HTNALAEURDDALRGHLRQVSLRHRNRLRALPTLFRNLSSL 389

Db 402 NANEISCTIRKDAFRDLHLSLSLLSYDNNISQSLANGTFDANKSM 444

Wed Dec 1 15:24:46 2004

us-09-775-803-12_copy_1_389.rpr

Page 8

Search completed: November 29, 2004, 13:11:29
Job time : 27 secs